Table of Contents

Executive Summary

During fiscal year 1998, UI agencies made over 10.6 million determinations of monetary eligibility for UI benefits, over 3.4 million separation determinations, and over 4.3 million nonseparation determinations (see Chapter 1 for details). The Denied Claims Accuracy Pilot Project was an operational pilot conducted during 1997-98 in five states with the purpose of gathering the information that will be needed to guide a program of measuring Denied Claims Accuracy in all 53 UI jurisdictions. This report describes and presents the main findings from the pilot project.

Design and Operations of the Pilot Project. Each of five pilot states that agreed to participate drew random samples of (roughly) 200 monetary denials, 200 separation denials, and 200 nonseparation denials and subjected them to intensive investigation (along the lines of the existing core Benefits Quality Control program) in order to determine their accuracy. Monetary denials were investigated using the BQC approach, which involved a review of all pertinent agency records, an interview with the claimant, and contacts with base-period employers to ascertain the correct wages, hours of work, weeks of work, etc., as prescribed by State law. Nonmonetary denials (both separations and nonseparations) were reviewed twice: first, under a BQC-type review involving a claimant interview and appropriate contacts with employers and/or third parties to determine the eligibility decision that would accord with a fully-informed application of State law and policy; and second, under a briefer review of applicable agency data alone, which resulted in a rating of the denial according to the existing Quality Performance Indicator (QPI) instrument. (Chapter 2 presents further details of the pilot's design.)

Each pilot state received resources for two DCA Pilot Project investigators. Sampling began in week 36 of 1997 (early September 1997) and continued until between week 33 and week 36 of 1998 (August through early September 1998), depending on the pilot state. Representatives of the National Office, Regional Offices, and the contractor made site visits to four of the five pilot states. Findings from the site visits include the following:

- The pilot states reported that 3 staff years would be needed to handle a case load of 200 of each type of denial (600 total).
- The pilot states reported greater difficulty in obtaining information from claimants during DCA investigations than they experience during core BQC investigations.
- The pilot states reported that difficulty in obtaining responses from claimants caused delays in completing cases. However, the pilot states do favor timeliness standards for the DCA program.
- Most pilot states agree that the sampling of monetary

denials should be delayed for two weeks (ten work days) from the date that the claim was filed, in order to avoid sampling claims that were initially denied but that will be redetermined based on confirmation of additional wages.

- The pilot states reported no greater difficulties with investigations involving claims than with those involving intrastate claims.
- Staffs of the pilot states generally indicated that they prefer intensive investigation of denied claims (as in the DCA Pilot Project) to the Quality Performance Indicator rating of nonmonetary denials
- The pilot states reported only minor (and easily resolved) start-up problems in implementing the COBOL sampling program and getting the DCA Pilot Project running.
- Most of the pilot states had suggestions for changes in the Data Collection Instrument. (See Chapter 3 and section 7.3.)
- Staffs in some of the pilot states indicated that training in the BQC methodology in general and in nonmonetary determinations in particular would be beneficial.

See Chapter 3 for further information on the pilot's operation and for the detailed findings from the site visits.

Summary of Main Quantitative Findings of the Pilot. Table ES-1 summarizes the essential findings of the 1997-98 DCA Pilot Project regarding the accuracy of denied claims. (See Chapter 4 for a full discussion of the quantitative findings of the pilot.) Column (1) shows the denied claim error rates unadjusted for the effects of appeals, redeterminations, or agency actions to resolve issues. Column (2) shows the error rates adjusted for the effects of appeals and redeterminations, column (3) shows the error rates adjusted for the effects agency actions to resolve issues, and column (4) shows the error rates adjusted for the effects of appeals, redeterminations, and agency actions to resolve issues.

Column (1) of Table ES-1 shows that, in the five pilot states together, the unadjusted monetary denial error rate averaged 16 percent, the separation denial error rate averaged 8.7 percent, and the nonseparation denial error rate averaged 15 percent. Both unadjusted and adjusted error rates tend to be lower for separation denials than for monetary and nonseparation denials. (Because of substantial differences among the five pilot states in laws, policies, and procedures, no conclusions about the quality of state administration can or should be drawn from interstate comparisons of the error rates.)

Column (2) of Table ES-1 shows that appeals and redeterminations had no impact on monetary denial error rates. Appeals and redeterminations did, however, reduce the

separation denial error rate in the five pilot states together by about 2 percentage points (or 22 percent), and reduced the nonseparation denial error rate in the five pilot states by about 1 percentage point (or 6 percent). Appeals and redeterminations, then, are more effective in reversing erroneous nonmonetary denials than in reversing erroneous monetary denials.

Column (3) of Table ES-1 shows that the State UI agencies, through their own actions, reduced the monetary denial error rate in the five pilot states together by nearly 5 percentage points (or nearly 31 percent). The agencies' actions also reduced the separation and nonseparation denial error rates in the five pilot states, but by only about 1 percentage point in each case. Agencies' actions, then, are more effective in reversing erroneous monetary denials than in reversing erroneous nonmonetary denials.

Table ES-1 also displays summary results of the 1987 denied claims pilot. The unadjusted error rates in the 1997-98 DCA Pilot Project are generally lower than those in the earlier pilot (except for nonseparation error rates, which are about the same in both pilots). In fact, the unadjusted error rates in the 1997-98 pilot are roughly comparable to the error rates after adjusting for appeals and redeterminations in the earlier pilot (again, with the exception of the nonseparation error rates). This suggests that the determinations process, before any self-correction, has improved over the past decade. However, appeals and redeterminations now do less to reduce denial errors than they did a decade ago. (It appears impossible to adjust the error rates from the 1987 pilot so as to take account of agency resolution of issues. Accordingly, columns (3) and (4) of Table 7-1 have no data on the 1987 pilot.)

A main objective of the DCA Pilot Project was to compare the results of comprehensive field investigations with the QPI assessment of the quality of the determinations process. Table ES-2 shows cross-tabulations of the accuracy of separation and nonseparation denials by whether the denial determination passed or failed the QPI review. The results suggest that the correlation between QPI and DCA is highly imperfect: In only one case is the coefficient of correlation between the DCA assessment and the QPI assessment greater than 0.3. In all five pilot states and for both separation and nonseparation denials, a significant percentage of denials that were determined improper by DCA passed QPI, and a significant percentage of denials that were determined proper by DCA failed QPI. Moreover, the QPI suggests that the determinations process is less accurate than comprehensive field audits show. The conclusion is that the QPI is a very noisy predictor of the accuracy of denials.

Chapter 5 provides a brief discussion of the difficulties of estimating the dollar impact of erroneous denials. That discussion concludes that there are both conceptual and practical problems with estimating the dollar impact of erroneous denials. However, with further research, it may be feasible to overcome the problems and produce estimates of

the dollar impact of erroneous denials that could be reasonable guides for policy.

Recommendations. Chapter 7 summarizes the recommendations stemming from both the quantitative findings reported in Chapter 4 and the process of visiting the pilot states The main recommendations are as follows:

- It appears that, once allowance is made for supervisory review, about 2 to 3 staff years, not 2, are needed to investigate 600 denials in a year. The UI Performs budget currently includes only 2 full-time equivalent positions per state for denials. Accordingly, we recommend maintaining staffing at this level and making the denials annual sample size 150 for each type of denial.
- Four of the pilot states completed only between 43 and 79 percent of interviews in monetary investigations, and only between 68 and 83 percent of claimant interviews in nonmonetary investigations. We expect that interview completion will be a continuing problem in conducting investigations of denials. Accordingly, the experience of states and investigators with the best records should be drawn on and assembled in a handbook to provide advice and techniques for increasing contact rates. Standards for "reasonable attempts and effort" in attempting to contact claimants should be produced, consistent with the BTQ "reasonable attempt" standards.
- States reported no greater difficulty (or cost) in investigating and assessing the accuracy of interstate denials than intrastate cases. Accordingly, interstate cases should be included in the national program.
- For the nationwide program, sampling of monetary denials should be delayed for up to 14 days so that wage-request processes are completed and most of the denials in the sample are "true" denials. Nonmonetary denials should be sampled and assigned for investigation at the end of the week in which they occurred, just as in the pilot.
- The pilot showed little correlation between the results of the DCA review of denied nonmonetary claims and the QPI review of the quality of nonmonetary decisions. The clear conclusion is that the accuracy of nonmonetary denials is best estimated using the BQC reassessment and reverification method.
- States should be given the same flexibility to select among the various review techniques as they have for paid claim accuracy.
- The pilot states found that the DCA review process worked best when all Quality Control staff shared denials (DCA) and payment (core BQC) cases. Accordingly, we recommend that states be given flexibility in assigning staff to DCA cases.

- The DCA Pilot Project had no standards for timely completion of cases. We recommend that time lapse standards be established for denials in the nationwide program, with the understanding that these standards may differ from the BAM standards.
- The pilot was intended to identify data elements which should be captured for denials investigations, or for which definition should be modified. The pilot states identified several elements in the Denied Claims DCI for which edits needed to be modified or which could be pre-filled or "stamped" in case they were inapplicable in the state. The pilot states recommended eliminating all elements designed to capture dollar impacts of denials. A workgroup of State and Regional staff has met to review all other elements of the Data Collection Instrument to be used in both Denials and Benefits Quality Control and to develop a revised and integrated DCI.

Table ES-1
Summary of Denial Case Error Rates by Type, Unadjusted and Adjusted for Appeals, Redeterminations, and Agency Resolution, 1997-98 and 1987 Denials Pilot Projects (percentage erroneous denials for all pilot states with five-state range in parentheses)

		Error Rate (%) Adjusted for:		
	(1)	(2)	(3)	(4)
	Error rate (%)	Appeals & re-	Agency	
Type of Denial	unadjusted	determinations	resolution	All factors
1997-98 Pilot				
Monetary	16.0 (10 - 23)	16.0 (10 - 23)	11.2 (8 - 16)	11.2 (8 - 16)
Separation	8.7 (3 - 20)	6.8 (3 - 17)	8.0 (3 - 19)	6.4 (3 - 16)
Nonseparation	15.0 (7 - 22)	14.1 (6 - 20)	13.8 (6 - 18)	12.9 (6 - 17)
1987 Pilot				
Monetary	23 (10 - 36)	16 (7 - 33)	na	na
Separation	15 (5 - 29)	9 (2 - 25)	na	na
Nonseparation	14 (7 - 23)	11 (6 - 21)	na	na

Source: Tables 1-5, 4-10A, 4-10B, and 4-10C.

Table ES-2
Crosstabulations of Separation and Nonseparation Denial Accuracy by Modified QPI Pass/Fail, All Five Pilot States
(frequencies with row percentages)

A. Separation Denial Accuracy by Modified QPI Pass/Fail

	QPI score		
DCA finding	fail	pass	Total
Improper denial	54 (67)	27 (33)	81 (100)
Proper denial	233 (28)	603 (72)	836 (100)
Total	287 (31)	630 (69)	917 (100)

B. Nonseparation Denial Accuracy by Modified QPI Pass/Fail

	QPI score		
DCA finding	fail	pass	Total
Improper denial	78 (55)	65 (45)	143 (100)
Proper denial	161 (21)	607 (79)	768 (100)
Total	239 (26)	672 (74)	911 (100)

Source: Tables 4-14 and 4-15.

List of Tables

Denied Claims Accuracy Pilot Project

Final Report

May 1999

1. Background

1.1. Development of the Denied Claims Accuracy Pilot Project

In 1984, the Unemployment Insurance Service of the U.S. Department of Labor developed a Benefits Quality Control (BQC) program. (This program is currently known as Benefits Accuracy Measurement; however, throughout this report, the original designation of "BQC" will be used.) In October, 1987, the "core" BQC program, which measures payment errors (particularly overpayments) became mandatory in 52 of the 53 Unemployment Insurance jurisdictions (all but the Virgin Islands). Originally, the Unemployment Insurance Service envisioned expanding the BQC program – after appropriate pilot testing – into aspects of the UI program other than the accuracy of intrastate paid claims. In particular, the agenda for quality control efforts included pilot tests for the accuracy of benefit denials, the accuracy of interstate payments, the accuracy of tax collections, and the use of telephone, fax, and the mail to perform quality control.

The General Accounting Office, the Office of Inspector General, organized labor, and advocates of claimants strongly urged the Department to expand BQC and implement a program that would estimate the accuracy of UI claims that were denied – in other words,

a "Denials Quality Control" or "Denied Claims Accuracy" subprogram. This was viewed as important in order to maintain balance in the way the Department treats claimant issues. During fiscal year 1987, in a DOL-State pilot, five states investigated the accuracy of claims that were formally (that is, officially) denied for monetary reasons, reasons of separation, and reasons of weekly eligibility.

Although the 1987 pilot suggested that a substantial number of denials were erroneous and were not corrected by the workings of the redetermination and appeal processes, other priorities and resource limitations precluded incorporating denials into the framework of the BQC program. Rather than proceed with measuring the accuracy of denied claims, the Unemployment Insurance Service conducted the Revenue Quality Control and Performance Measurement Review initiatives. In 1993, however, the Department of Labor agreed with the Vice President's National Performance Review that the BQC program needed to be reexamined. According to the National Performance Review issue paper, the Department, as part of its reexamination of BQC, needed to consider whether BQC should "continue to keep its existing focus on paid claims ... or include measurement of decisions to ... deny claims."

The BQC reexamination was undertaken by a joint Federal-State Performance Enhancement Work Group and occurred within the context of an overall review and restructuring of UI performance measurement. The Performance Enhancement Work Group first proposed a new approach to unifying all UI performance measurement, and within this system it proposed a benefit accuracy measurement program considerably

smaller than BQC that offered states more flexibility to use telephone, fax, and mail to verify information. The Performance Enhancement Work Group also recommended that the accuracy of denials be assessed.

There has been continuing interest in assessing the accuracy of denied claims for other reasons as well. In its 1994 performance review of the Unemployment Insurance Service's Consolidated Financial Statements, the Department of Labor's Office of Inspector General recommended that the Unemployment Insurance Service "initiate quality control programs to measure the accuracy of denied initial claim determinations which should be quantified and reported as underpayments in the financial overview." In a March 23, 1995, letter to Deputy Secretary of Labor Thomas Glynn, the United Auto Workers recommended that "benefits quality control should be modified to measure and report wrongful denials of UI benefits." Also, in a March 1995 meeting with Department of Labor officials, representatives of the National Employment Law Project said they considered the lack of a means for measuring the accuracy of denials to be a definite deficiency.

Because of the time that had elapsed since the 1987 pilot, given the need to address various questions that were left unresolved after the earlier pilot, and at the urging of the Deputy Secretary of Labor, the Unemployment Insurance Service decided to conduct another pilot in order to test the measurement of denials accuracy before implementing such a program nationwide.

This report describes that pilot – known as the Denied Claims Accuracy Pilot Project. The plan of the report is as follows. The remainder of this chapter presents the

context of the Denied Claims Accuracy Pilot Project by discussing the processes by which eligibility for UI benefits is determined and the ways in which claimants may be denied UI benefits. Chapter 1 also reviews basic concepts of accuracy measurement and reviews the conduct and results of the 1987 Denials Pilot Project.

Chapter 2 describes the Denied Claims Accuracy Pilot Project, its purposes, and its design. Chapter 3 presents findings and recommendations from site visits that were made by representatives of the UI Service (both the National Office and Regional Offices) and the contractor (PRAMM Consulting Group, Inc.). Chapter 4 presents a quantitative analysis of the findings of the DCA Pilot Project. Chapter 5 discusses whether it might be possible to estimate the dollar implications of erroneous denials – that is, to estimate how much would have been paid in benefits if the erroneous denials had been determined correctly. Chapter 6 is a brief discussion of the costs of conducting a DCA program. Chapter 7 summarizes the report and presents conclusions and recommendations for implementation of the DCA program on a national basis.

1.2. UI Benefits Activity and Denials

There are three ways in which UI benefits may be denied to workers. First, a potential claimant might conclude that he or she is not likely to be eligible for benefits and hence never contact the UI agency. This might be called a "self denial." Second, a claimant might make an informal inquiry, usually about monetary eligibility (see below), and be advised that he or she is not eligible or not likely to be eligible. As a result, the

potential claimant never files a formal initial claim. This case has frequently been termed an "informal" or "counter" denial. Third, if a formal initial claim is filed, UI eligibility is tested, more or less sequentially, at three levels, and there is the potential for a formal denial at each level. Note that (except in section 3.4) this report focuses only on formal denials. What follows is a description of the three formal criteria by which eligibility for benefits are tested.

1.2.1. Monetary Criteria. In order to be eligible for UI benefits, claimants must have earned enough wages or worked enough weeks (or both) during a "base period," which is usually the first four of the past five completed calendar quarters before the initial benefit claim was filed. If the work and pay records available to the UI agency indicate that the state's monetary conditions are not satisfied, a "monetary denial" of benefits remains in effect until enough additional wages to qualify are earned. (In some cases, the records available to the agency are incomplete, and the denial remains in effect until the missing work and pay records appear or are identified).

As can be seen in Table 1-1 (top panel), in fiscal year 1998, about 10.6 million determinations of monetary eligibility were made. Although not shown in the table, about 90 percent of these determinations were made on initial claims (officially designated "new initial claims"). The remainder were made on transitional claims – claims filed by claimants who remain unemployed when their original benefit year expires and who attempt to establish a new benefit year on the basis of earnings during the lag quarter. The lag quarter is the quarter immediately preceding the quarter in which the initial claim was filed

and is not part of the traditional base period. (To be eligible for benefits, such a claimant must generally have had some earnings subsequent to the lag quarter as well.) In 1998, 1.2 million, or 11 percent, of the monetary determinations resulted in a determination of monetary ineligibility.

1.2.2. Separation Criteria. To be eligible for UI benefits, a claimant must also have been separated from employment due to lack of work and through no fault of his or her own. This means that workers who quit a job voluntarily are generally ineligible for UI benefits, although a worker who had "good cause" for quitting may be eligible for benefits (the definition of "good cause" varies from state to state, as discussed in section 2.3 below). It also means that a worker who is discharged for misconduct will be ineligible for UI benefits. Depending on the state and the specific reason for failing the separation criterion, a separation denial can make a claimant ineligible for periods ranging from several weeks to indefinitely. In the case of an indefinite separation denial, a worker will generally be required to find reemployment and to earn "requalifying wages" before being eligible for benefits.

Separation determinations can be made on initial claims and reopened claims – that is, claims that have the potential to start a new series of UI benefit payments. (Reopened claims – officially, "new additional claims" – are claims made during an existing benefit year after an interruption in benefit receipt.) The second panel of Table 1-1 shows that, in 1998, there were about 15.4 million initial and reopened claims. UI agencies raised and adjudicated separation issues with respect to 3.4 million (or 22 percent) of these initial and

reopened claims. In turn, 55 percent of the adjudications resulted in denial of benefits. Accordingly, 12 percent of all initial and reopened UI claims resulted in denial of benefits based on conditions of separation.

1.2.3. Nonseparation Criteria. Finally, a claimant who is monetarily eligible for UI benefits and who separated from employment for acceptable reasons must be able to work, available for work, and actively seeking work if required by State law or policy. An otherwise eligible claimant who is not "able, available, and actively seeking" work is denied UI benefits for reasons of "nonseparation." A nonseparation denial usually involves loss of eligibility for the claimed week, although a worker who refuses to accept suitable work be denied benefits for several weeks.

In 1998, there were about 131.6 million weekly claimant contacts that could have resulted in a nonseparation determination of weekly eligibility (see the bottom panel of Table 1-1). Of these, State agencies identified and formally adjudicated about 4.3 million issues (that is, 3.3 percent of the contact weeks were adjudicated). The adjudications resulted in about 2.4 million denials. Hence, about 1.8 percent of all weekly claimant contacts resulted in denial based on nonseparation conditions.

An issue of nomenclature should be raised at this point. Separation and nonseparation criteria for UI eligibility are sometimes referred to together as "nonmonetary" criteria. As a result, denials that occur because a worker's conditions of separation were unacceptable are sometimes referred to not simply as "separation" denials but rather as "nonmonetary-separation" or "nonmon-sep" denials. Similarly, denials

that occur because a worker was not "able, available, and actively seeking" work are sometimes referred to not simply as "nonseparation" denials but rather as "nonmonetary-nonseparation" or "nonmon-nonsep" denials. Throughout this report, we use the shorter designations – "separation" denials and "nonseparation" denials – to refer to the two types of nonmonetary denial, understanding that both separation and nonseparation denials result from applying nonmonetary criteria for eligibility.

1.3. Administrative Data on UI Determinations and Denials

As discussed in the previous section, State UI agencies make millions of determinations each year regarding monetary and nonmonetary eligibility of claimants. This section provides further detail on UI benefit activities and denials. In particular, we present information reported by the states on monetary determinations (Table 1-2), separation determinations (Table 1-3), and nonseparations (Table 1-4). Data are shown separately for each of the five states that participated in the Denied Claim Accuracy Pilot Project (Nebraska, New Jersey, South Carolina, West Virginia, and Wisconsin), along with national totals. Our purpose is to compare UI benefit activities and denials in each of the five pilot states (and in the five pilot states combined) with the national experience.

1.3.1. Monetary Issues. Reports on monetary determinations and nonmonetary determinations (ETA 218 and ETA 207, respectively) are submitted quarterly by the states. The time period in these reports that most closely approximates the period when the pilots were actively collecting data was from October 1997 through September 1998, i.e., federal

fiscal year 1998. Although actual data collection in the pilot states typically commenced slightly earlier, in mid-September 1997, and ended in late-August or early-September 1998, the timing of the data collection closely matched fiscal year 1998.

Nationwide during fiscal year 1998, over 11 percent of monetary determinations concluded that the claimant had insufficient wage credits for monetary eligibility. Table 1-2 shows there were 10.6 million monetary determinations and 1.2 million denials for a monetary denial rate of 0.114. The five pilot states combined made a total of 0.94 million monetary determinations of which 0.10 million had insufficient wage credits. Hence, the average monetary denial rate in the pilot states – 0.106 – was only slightly lower than the national average of 0.114.

When the individual states are examined in Table 1-2, however, the range of monetary denial rates is observed to be quite wide. Nebraska's denial rate of 0.029 was only about one quarter of the national average. West Virginia and Wisconsin also had denial rates well below the national average – 0.065 and 0.069 respectively. In contrast, New Jersey and South Carolina had monetary denial rates that exceeded the national average during this period – 0.142 and 0.131 respectively.

Inferences about incorrect monetary denials cannot be made from the program data in Table 1-2. However, if the case error rates from the denials pilot are representative of national experience, the overall rate of erroneous monetary denials probably falls in the 0.13-0.19 range. The average case error rate was 0.16 across the five pilot states (144 of 901 monetary determinations). Applying the average case error rate of 0.16 to the 1.2

million determinations of insufficient wage credits would imply that more than 190,000 erroneous monetary denials (out of 10.6 million determinations) were made throughout the United States during fiscal year 1998. This is nearly 2 percent of all monetary determinations.

1.3.2. Separation Issues. Table 1-3 summarizes separation denials in fiscal year 1998. Unlike monetary determinations, which occur only when a claimant establishes a new benefit year, separation determinations are associated with both initial and reopened claims, the latter occurring after an interruption in benefit receipt during an established benefit year. Panel A of Table 1-3 displays initial and reopened claims for each state, for the five pilot states combined and the U.S. totals for fiscal year 1999. (In fact, what is shown as "initial and reopened claims" in Table 1-3 is the sum of (a) new initial claims times the fraction monetarily eligible, (b) interstate liable claims, and (c) reopened claims. The figures come from the ETA 218 and ETA 207 reports.) Two issues, voluntary quits and misconduct, account for nearly all separation determinations in the states. Therefore, Table 1-3 displays totals plus detail for these two issues. Panel B displays two types of separation activity measure: determination rates (determinations as a proportion of initial and reopened UI claims) and denial rates (denials as a proportion of initial and reopened UI claims). Both the determination rate and the denial rate affect the total number of separation denials. That is, a high volume of denials can be obtained by various combinations of a high determination rate and a high denial rate per determination.

The first feature that is clear in Table 1-3 is Nebraska’s unusually high

determination rates for both voluntary quits and misconduct. In fact, Nebraska has the highest separation determination rate of the 53 UI jurisdictions because it adjudicates every separation that may be relevant to monetary eligibility. Nationwide, the average separation determination rate was 0.22. In the other four pilot states, the separation determination rate ranged between 0.134 (South Carolina) and 0.238 (Wisconsin). In Nebraska, however, the proportion was 0.947; that is, there were nearly as many separation determinations as there were initial and reopened UI claims. While determinations on misconduct issues were also frequent in Nebraska (more than twice the national average), the determination rate for voluntary quits was more than six times the national average (0.679 versus 0.096).

As noted, separation determination rates in the pilot states other than Nebraska are not far from the national averages. Nationally (and in the four pilot states other than Nebraska) there is a somewhat higher determination rate for misconduct issues than for voluntary quits. Nationwide, the average was 0.118 for misconduct and 0.096 for voluntary quits in fiscal year 1998. For all five pilot states (including Nebraska), the corresponding averages were 0.121 and 0.102, both just above the national average.

Separation denial rates are noticeably higher for voluntary quits than for misconduct. The national average denial rates in Table 1-3 were respectively 0.737 and 0.407. Compared with these averages, the averages from the pilot states were slightly higher: 0.802 for voluntary quits and 0.445 for misconduct. Nebraska, New Jersey, South Carolina and West Virginia had above average denial rates for voluntary quits (all above

0.850) while South Carolina also had a very high denial rate for misconduct (0.758). Because denial rates are much higher for voluntary quits, total denials resulting from this issue exceed misconduct denials (1.09 million versus 0.74 million) despite the fact that states made more misconduct determinations than voluntary quit determinations (1.83 million versus 1.48 million).

When the number of separation denials is related back to initial and reopened UI claims, Nebraska again is the outlier. Nationwide and in the four pilot states other than Nebraska, separation denials as a proportion of initial and reopened claims ranged between 0.107 and 0.120. In Nebraska, however, the corresponding proportion was 0.785 (these figures are not displayed in Table 1-3). The national proportion of 0.120 means there were 1.86 million denials due to separation issues in fiscal year 1999.

Data from the five pilot states indicate that 8.7 percent of denials on separation issues were improper (87 of 1,006). Applying this case error rate to the 1.86 million separation denials of fiscal year 1998 would imply that, nationwide, there were over 160,000 improper denials on separation issues for this period.

1.3.3. Nonseparation Issues. Table 1-4 summarizes nonseparation determinations with totals and details for four different types of nonseparation issues. Because a nonmonetary determination could potentially occur for any week in which benefits are claimed, we sum continued weeks claimed and initial claims and refer to this sum as claimant contacts (that is, the sum of weeks claimed). We then define the nonseparation determination rate as the number of nonseparation determinations per 10 claimant

contacts. (Since nonseparation determinations are relatively rare in the case of initial claims, it can be argued that initial claims should be excluded from the measure of claimant contacts. Excluding initial claims from claimant contacts would reduce claimant contacts nationally from 131.6 million to about 115 million, and would raise the national nonseparation determination rate from .329 [that is, 3.3 percent] to 0.374 [3.7 percent]. Thus, excluding initial claims from claimant contacts would change only slightly the nonseparation determination rates shown in Table 1-4.) The nonseparation denial rate, in turn, is the number of nonseparation denials per nonseparation determination. In fiscal year 1998, there were almost 132 million claimant contacts in all UI programs. For this same period there were more than 4.3 million determinations on nonseparation issues and almost 2.4 million nonseparation denials.

The meaning of nonseparation denials should be stressed. Most refer only to the week of the determination and not to the entire remaining period of potential eligibility. This contrasts with separation determinations, where denials usually apply to the entire spell under consideration. It follows that a denial on, say, an able and available issue will pertain just to the week in which the issue has been raised. Denials on nonseparation issues typically have less severe consequences for the claimant than denials on separation issues. (This argument does not apply to all nonseparation denials. If there is a refusal of suitable work – a less common nonseparation issue – the denial can apply to the entire remaining period of entitlement.)

The four nonseparation issues identified in Table 1-4 account for most

nonseparation determinations. Combined, they totaled 4.12 million, or 95 percent of all determinations on nonseparation issues. The issue "other" is a catchall that includes alien applicants, athletes, and covered employees of school systems.

In the pilot states taken together, these four issues are generally less common than in the rest of the United States. Whereas the five pilot states accounted for 0.085 of claimant contacts, they accounted for lower shares of determinations on able and available, disqualifying (and deductible) income, and "other" issues. They had an above average representation only among nonseparation determinations related to reporting requirements (0.096). They had especially low representation among determinations relating to disqualifying income (0.024).

Within the five pilot states, overall determination rates are above average in Nebraska and Wisconsin. In the other three states, the overall determination rates are half or less of the national average. Nebraska has high determination rates for reporting requirements and for disqualifying income. Wisconsin has high determination rates for reporting requirements and "other" nonseparation issues.

In the United States overall, nonseparation denials occur in over half (0.545) of the determinations that are made. Overall and for three of the detailed determination categories, the average denial rate in the pilot states exceeded the national average. Denial rates are especially high in Nebraska. West Virginia was the only pilot state where the average nonseparation denial rate was less than the national average.

Results from the pilot project suggest the case error rate data on nonseparation

denials to be 0.150. In fiscal year 1998, there were a total of 2.36 million denials on nonseparation issues. Applying the case error rate from the pilot project to these national totals would imply an estimated total of 350,000 erroneous denials on nonseparation issues during fiscal year 1998.

Combining the rough estimates made here from all three types of erroneous denials would yield a total estimate of about 680,000 erroneous denials of UI claims during fiscal year 1998. This is the sum of 170,000 involving monetary determinations, 160,000 involving separation issues, and 350,000 involving nonseparation issues. This total gives an indication of the large potential for erroneous decisions affecting worker entitlements to UI benefits.

1.4. Measurement of the Accuracy of UI Denials

To a large extent, self denials and informal denials are beyond the reach of the UI system (but see section 3.4 below). Accordingly, the Department of Labor has regarded the extent and correctness of self denials and informal denials to be a matter for research, not for operational measurement. At present, the UI system does not assess the accuracy, per se, of denied claims, although as noted in section 1.1, the Department has recognized the importance and feasibility of measuring the accuracy of denied claims. The quality of most nonmonetary determinations, including denials, is regularly appraised through the Quality Performance Indicator Assessment (discussed next), but monetary denials are not appraised in any way. At present, the UI system has two mechanisms that could be used

to measure the accuracy of all or some denied claims.

1.4.1. The Quality Performance Indicator (QPI) Assessment of Nonmonetary Determinations. Since the 1970s, Ul agencies have assessed the "quality" of nonmonetary adjudications by reviewing a sample of adjudications using the Quality Performance Indicator (QPI), a part of the Quality Appraisal program. About 60 percent of the determinations reviewed under QPI during fiscal year 1998 were denials, 34 percent were eligible claims, and information was not available for 6 percent of the cases. The QPI review basically assesses the quality of the processes of applying the separation and nonseparation criteria. The QPI was not designed to determine whether the process reached the fully correct decision given State law and policy, nor are the results tabulated to determine the share that were accurate. The QPI notion of accuracy is limited to determining whether the state (a) took proper steps to obtain information, (b) offered the opportunity for rebuttal, (c) correctly identified an issue, and (d) correctly applied appropriate law. The accuracy of the information used to make the determination is not considered in the QPI process. Neither is missing information due to late response or no response from one or another party considered.

1.4.2. Benefits Quality Control (BQC). The BQC program, implemented under regulation in October 1987, investigates the accuracy of UI benefit payments and yields estimates of UI benefit overpayments. It does not measure the accuracy of denials. The BQC program samples only payments (specifically, weeks paid), and accurately estimates the number of overpayments and dollars overpaid. However, it underestimates

underpayments because it samples only weeks in which UI benefits were paid – it does not sample weeks that should have been paid but were not. By thoroughly reviewing and field-verifying all information relating to the accuracy of payment decisions, BQC investigators determine whether UI staff had the necessary and correct information to work with and whether information obtained by UI staff was used properly. Accordingly, BQC provides an assessment of whether decisions were made with full information and in accordance with law and policy. Moreover, BQC assesses both monetary determinations and nonmonetary (that is, separation and nonseparation) determinations. (Recall that QPI applies only to nonmonetary determinations.) In the case of separation and nonseparation determinations, the BQC review ascertains whether all issues were correctly identified by UI staff and whether identified issues were correctly handled. It also ensures that all appropriate information is marshaled for the adjudications. Hence, the BQC review often obtains information not received previously by the agency or not submitted timely to the agency.

The "BQC methodology" involves drawing samples of paid claims. As already outlined, the decision to pay a weekly benefit follows a sequence of three decisions. Reviewing the correctness of the payment requires BQC to verify all prior decisions on which the decision to pay the particular week claimed depends. As a result, BQC investigates "up the chain" of UI decisions. In 1994, BQC reviewed (through its samples) the universe of 120.4 million weeks paid. In addition, it reviewed the 10.4 million monetary decisions and 13.3 million separation decisions that had to be positive in order for those

120.4 million paid weeks to be correct. In 1994, the universe of decisions investigated by BQC sampling amounted to approximately 97 percent of all decisions made in intrastate UI, UCFE, and UCX cases. (Interstate payments are not investigated in the BQC program.)

As noted, several groups have stressed the need for balance in BQC by using its approach to assess the accuracy of denied claims as well as paid claims. This concern for balance led to the 1987 pilot testing of several approaches (including the existing BQC approach outlined here) soon after payments (or "core" BQC) became operational.

1.5. The 1987 Denials Pilot Project

During fiscal year 1987, with the technical assistance of Applied Management Sciences, Inc., five states pilot tested three different approaches to assessing the accuracy of formally denied claims using the BQC field-verification approach. The five states were Louisiana, Pennsylvania, Iowa, South Carolina and Washington. (This section relies on Belle and Casey 1988.)

1.5.1. Design of the 1987 Pilot. Three approaches were taken to selecting samples of cases to be investigated in the five states:

- Approach 1: Separate cross-sectional weekly samples of denied claims, with BQC maintained. One state defined separate universes of the three types of denial decisions (monetary, separation, and nonseparation). It drew and investigated weekly samples of each. This state's BQC program remained unchanged and continued to measure the accuracy of paid claims in its usual way. This approach was adopted for the 1997-98 DCA Pilot Project.

- Approach 2: Separate cross-sectional weekly samples of positive and negative monetary, separation, and nonseparation decisions. In one state, the process of determining eligibility was disaggregated into its three stages or levels of

decisions: monetary, separation, and weekly nonseparation eligibility. At each level, samples of both positive and negative decisions were drawn and investigated (as in BQC) to determine their accuracy. This model offered an alternative to the existing BQC design.

- Approach 3: Longitudinal tracking. Three states (Iowa, South Carolina, and Washington) tested a longitudinal approach, which, like Approach 2, involved a redesign of BQC. Each week, a sample of initial claims was drawn and added to a "tracking file." The experience of this cohort was monitored; all denials were investigated as they occurred, beginning with monetary denials. In addition, a sample of the weeks paid to the claimants remaining in each cohort was investigated each week, providing an alternative to BQC's method of selecting payments for review.

Although the five states used three different methods for selecting samples of denials, once the samples were selected, all investigated the cases in the same way, following as much as possible the in-person BQC methodology that was current at the time (there is now greater reliance on telephone, fax, and mail in conducting investigations). In addition to reviewing all pertinent agency records, each investigation involved an interview with the claimant and contacts with the parties necessary to ascertain the facts on which the denial decision was based.

1.5.2. Summary of Findings. The 1987 pilot investigated the correctness of denied claims before redetermination or appeal, and also noted whether those denials were ultimately reversed by the appeal and redetermination process. Table 1-5 displays the percentages of the decisions to deny benefits that the 1987 pilot found to be in error (that is, should have approved as eligible UI claims) both before and after appeal and redetermination.

The 1987 pilot suggested that, on average, existing appeal or redetermination

processes reverse one-fifth to two-fifths of erroneous denials (depending on the state). Nevertheless, between one-tenth and one-sixth of denials remain erroneous.

The pilot itself yielded only case error rates. Because claimants who are denied benefits have no observable claim experience, the dollar impact of erroneous denials – benefits lost by claimants – can be estimated or imputed only through some form of statistical modeling (see Chapter 5 for further discussion). The contractor did attempt this in the 1987 pilot, but the number of assumptions required to make such projections makes them of questionable value for guiding decisions.

In addition to investigating the correctness of denials, states in the 1987 pilot applied the usual BQC action, cause, and responsibility codes to denial errors. As Table 1-6 shows, the UI agencies were attributed either partial or total responsibility (the contractor's report did not differentiate between these) for about one-quarter of the erroneous monetary denials and for about three-quarters of erroneous nonmonetary (separation and nonseparation) denials.

Finally, the 1987 pilot characterized the agencies' actions on erroneous denials in one of five ways: (a) the source of the issue was undetectable with existing procedures, (b) the issue was already being resolved by the agency, (c) the wrong action was taken by the agency, (d) there was insufficient follow-through by the agency, and (e) the issue was not detected because the agency did not follow procedures. The percentages of erroneous denials for which each of these agency actions was taken are shown in Table 1-7.

The range associated with each of the actions taken is quite large, in part because the estimates are based on errors detected, the absolute number of which was frequently quite small. However, on average, Table 1-7 suggests that existing procedures were able to detect about three-quarters of all types of erroneous denials. In roughly one-quarter to one-half of all cases, the agency was already resolving the error. For nearly one-third of the erroneous separation and nonseparation denials, the agency had the necessary information but took the wrong action. (The agency took the wrong action in relatively few monetary cases.) For the remaining one-sixth (roughly) of all erroneous denials, the agency did not follow through on information it had or missed the issue because it did not follow its procedures.

2. Description of the Denied Claims Accuracy Pilot Project

2.1. Reasons for a New Pilot Project

Several of the factors impelling the Department of Labor to measure the accuracy of denied claims have been outlined in section 1.1. Based on the experience of the 1987 pilot and further discussion, the Department decided in 1995 to measure denied claims accuracy by drawing three separate (cross-sectional) samples of denial decisions on a weekly basis – one each for monetary, separation, and nonseparation denials – and investigating each decision using the BQC methodology. This was one of the three alternative sampling designs (Approach 1) tested in the 1987 pilot (see section 1.5 above), and it was chosen as the most satisfactory sampling design because, both conceptually and practically, it closely resembles the sampling used to conduct the existing BQC program. As a result, it involves no fundamental redesign of the sampling procedures used for BQC.

Rather that attempt immediate nationwide implementation, however, there were three reasons for conducting a second pilot. First, nearly ten years had elapsed since the 1987 pilot, and it was unknown whether conditions had changed substantially since then. For example, changes in determination rates or denial rates could change the resources needed by State agencies to measure denied claims accuracy. If so, then those changed conditions would need to be taken into account in the design of a nationwide program.

Second, even if conditions were similar, the experience of a few pilot states would be highly useful to the other states in implementing the denied claims accuracy program. Third, several questions were left unresolved after the 1987 pilot, and additional information was needed to address these questions and guide implementation of a national measurement effort. These questions are discussed next.

2.2. Questions to Be Addressed and Issues to Be Resolved

The 1997-98 Denied Claims Accuracy Pilot Project had the goal of addressing the following four principal questions.

1. What are the resource implications of investigating denials accuracy? Answering this question requires answers to the following:

- How large should be the samples of denied claims that are investigated? Larger samples yield more precise estimates but require more resources. There is a tradeoff between the precision of the estimates and the resources needed to obtain them.

- How long do investigations of denied claims take? The 1987 pilot concluded that, on average, a denials investigation according to the BQC in-person protocol required about 60 percent as much time as a payment investigation. Since that time, however, the BQC protocol has changed. In 1993, an "alternative methodology" was adopted, and in January 1997, the Department allowed states complete flexibility to use telephone, mail, and fax to carry out investigations. How long a denials investigation takes will depend importantly on how the investigations are conducted – using the present mix of in-person and phone, mail,and fax contacts; using all phone, mail, and fax; or relying more or less entirely on in-person methods. Another determinant will be the extent of the denials investigation itself.

- In addition to changes in the way BQC investigations are conducted (mentioned above), it has since been recommended that the BQC approach should be used only to investigate the accuracy of monetary denials. The rationale for this

recommendation stems from existence of the Quality Performance Indicator (QPI) review of nonmonetary determinations (both separation and nonseparation denials). This recommendation would have states use the QPI method and instrument to review samples of separation and nonseparation denials, rather than adopt the BQC approach (as has been done in the 1997-98 DCA Pilot). If this recommendation were accepted, it would have considerable resource implications because it involves only a review of existing records. It would not involve recontacting the parties to the adjudication, as does the BQC approach.

- Since the BQC approach is to be used, what should be the scope of the investigation for nonmonetary denials? That is, should a separation investigation focus only on the circumstances of the separation, or should it include a review of monetary eligibility as well? Should investigation of a nonseparation denial be limited to that particular issue, or should it include monetary and/or separation eligibility as well? (Under BQC, which was followed in general for the 1997-98 DCA Pilot Project, investigation of a separation denial does involve investigating prior monetary eligibility, and investigation of a nonseparation denial involves investigating prior monetary and separation eligibility. It should be noted, however, that under core BQC, a denial cannot be the starting point of an investigation; rather, a denial arises only when BQC reverses an original positive determination during an investigation.)

2. In sampling and in conducting the investigations, what allowances should be made for the normal processes of redeterminations and appeals? In other words, at what time should samples of denials be drawn, and at what time should they be investigated?

3. What modifications to the Data Collection Instrument that is used in BQC are needed for the Denied Claims Accuracy program? A basic assumption guiding the DCA Pilot Project was that the measures of interest are rates of correct and incorrect denial decisions of the three main types. (That is, the DCA Pilot Project did not attempt to estimate or project dollar impacts of erroneous denials.) In order to estimate erroneous denial rates, however, the states needed to determine what information would need to be gathered in the course of the denials investigation. Accordingly, decisions needed to be

made about which elements to include and which to drop from the existing BQC Data Collection Instrument. Decisions also needed to be made about adding any new elements. Only the experience of states that were actively investigating denials and estimating denial rates could guide these decisions.

4. What is the appropriate means, and level of detail, for obtaining information on the cost of investigating the various types of denial issues?

2.3. States in the Pilot and Their Characteristics

As already noted, the pilot involved five participating states – Nebraska, New Jersey, South Carolina, West Virginia, and Wisconsin. These states were selected from volunteers responding to a system-wide announcement that requested applicant states to indicate the following:

- the name and qualifications of their proposed candidate for pilot coordinator;
- their ability to identify the populations of monetary, separation, and nonseparation denial actions from automated records;
- their ability to build transaction files and implement COBOL programs, and maintain both;
- their interest and capability to write BQC-type (COBOL) sample selection programs;
- their commitment to making UI-qualified investigators available for the pilot; and
- whether their UI law has features that might affect the difficulty or extent of denials; for example, the existence of an alternative base-year law.

In selecting the five states, the Department of Labor attempted to achieve a reasonable balance among considerations of geography, state size, and features of the UI law.

Some general characteristics of the five pilot states are displayed in Table 2-1. The

pilot states represent a broad range of states by size of population and labor force, recent population growth rate, geographic size, population density, average pay, and labor market tightness (as measured by the unemployment rate). One of the pilot states, New Jersey, is among the most populous states (ranking ninth in population), and another, Wisconsin is a medium-sized state (ranking eighteenth). South Carolina's population has grown slightly faster than the U.S. average in the 1990s, whereas the other four states have grown more slowly than the U.S. average (with West Virginia and New Jersey being well below the U.S. average in their growth rates). Geographically, Nebraska is the largest of the pilot states and has very low population density, whereas New Jersey is the most densely populated state in the country. Except for New Jersey, which has relatively high average annual pay, the pilot states are near the U.S. average in annual pay. Finally, West Virginia's labor market is slack (an unemployment rate of 6.9 percent in 1997) compared with the national average, whereas Nebraska and Wisconsin have labor markets that are significantly tighter than the national average (unemployment rates of 2.6 and 3.7 percent).

For the purposes of the pilot, the diversity of UI eligibility criteria among the five pilot states is also important. Table 2-2 gives a summary of the monetary, separation, and nonseparation eligibility criteria in the five pilot states. The table and following discussion are based on the U.S. Department of Labor's Comparison of State Unemployment Insurance Laws (1997), Advisory Council on Unemployment Insurance (1995, Chapters 7 and 8), Nicholson (1997), Anderson (1997), and the laws of each of the five pilot states.

2.3.1. Monetary Eligibility Criteria. Table 2-2 characterizes the monetary eligibility

criteria of the states along four dimensions: (a) the type of formula used to determine eligibility, (b) the minimum base period earnings required for eligibility, (c) the type of earnings distribution requirement, and (d) whether there are any restrictions on the eligibility of seasonal workers.

All four types of formula for determining monetary eligibility are represented among the five pilot states. Nebraska and West Virginia are among the 6 states that use the so-called "flat" formula, under which a claimant must have earned some specified minimum dollar amount during the base period. (In addition, both require that base period earnings be earned in more than one quarter.) New Jersey requires that a claimant have worked at least 20 weeks (at specified minimum weekly earnings) during the base period. (New Jersey is one of 7 states that require a worker to have worked a minimum number of weeks or hours in the base period.) South Carolina is one of 24 states using a "multiple of high-quarter wages" formula. South Carolina's formula requires a worker to have earned at least $600 in the high quarter of the base period (that is, the quarter of the base period in which earnings were highest) and to have earned at least 1.5 times the high-quarter earnings (that is, at least $900) in the entire base period. Finally, Wisconsin is one of 14 states using a "multiple of weekly benefit amount" formula. Wisconsin' s formula requires first that a worker have earned at least $1,325 in the high quarter of the base period, which would qualify the claimant for a weekly benefit amount of $53. The worker must then have earned at least 30 times that calculated weekly benefit amount (that is, at least $1,590) during the entire base period to be monetarily eligible.

The second row under monetary eligibility criteria in Table 2-2 gives the minimum base period earnings that a claimant would need in order to qualify for benefits, as computed by the U.S. Department of Labor (Comparison of State Unemployment Insurance Laws, 1996, Table 301). The third row shows each state's distribution requirement. For example, Nebraska requires the claimant to have earned at least $400 outside the high quarter of the base period (the high quarter is the quarter of the base period in which earnings were highest). Similarly, the other five pilot states all have some requirement that the claimant's base period earnings were not concentrated in a single quarter. New Jersey, South Carolina, and West Virginia each require that a claimant have earnings in at least two quarters of the base period. Wisconsin requires the claimant to have earned at least 8 times his or her weekly benefit amount outside the high quarter of the base period.

Finally, the fourth row under monetary eligibility criteria in Table 2-2 shows that West Virginia and Wisconsin have special provisions that restrict the eligibility of seasonal workers. In general, wages earned by workers in an industry defined as seasonal can be used to establish UI eligibility only for unemployment during periods when the worker is usually employed in his or her seasonal job.

2.3.2. Separation Criteria. The middle panel of Table 2-2 summarizes the two main aspects of the separation criteria that a claimant must meet in each of the pilot states. The first is how a state handles the UI eligibility of a worker who voluntarily quit his or her job. Such a worker will be eligible for UI benefits only if he or she has "good cause" for quitting,

and the definition of good cause differs from state to state. Table 2-2 shows that in New Jersey, West Virginia, and Wisconsin, the definition of "good cause" is restricted to issues directly related to work or the employer, although West Virginia and Wisconsin define certain specific reasons for voluntary leaving that are permitted. (These include such reasons as illness of the claimant and leaving to accept another job that does not materialize. The number of such specific reasons is shown in the "inclusions" row of Table 2-2.) Nebraska and South Carolina, on the other hand, allow "good cause" to include good personal reasons as well as reasons that are directly related to employment.

Table 2-2 also shows the length of time a claimant who quits voluntarily will be disqualified from receiving UI benefits. In Nebraska, a worker who quits voluntarily is disqualified for 7 to 10 weeks. In the other pilot states, a worker who quits voluntarily is disqualified for the duration of his or her current spell of unemployment. In order to requalify for benefits, the worker must then earn some minimum amount, specified as a multiple of the weekly benefit amount in New Jersey, South Carolina, and Wisconsin. In West Virginia, the worker must work at least 30 days to requalify.

The second aspect of the separation criteria summarized in Table 2-2 pertains to misconduct. In most states, a worker who is discharged for misconduct will be ineligible for UI benefits for some specified number of weeks. In the pilot states, the duration of disqualification ranges from 5 weeks in New Jersey to up to 26 weeks in South Carolina. In addition, Table 2-2 shows that all of the pilot states except New Jersey reduce the benefits that a worker receives at the end of a disqualification period stemming from

misconduct. Finally, all of the pilot states except Wisconsin impose one or another additional penalty on a worker who was discharged for gross misconduct.

2.3.3. Nonseparation Criteria. Essentially, the nonseparation criteria for continuing eligibility specify that an individual must be able to work, available for work, and (usually) actively seeking work. The bottom panel of Table 2-2 summarizes the main features of the nonseparation criteria that a claimant must meet in each of the pilot states.

As noted by Anderson (1997) all the states require that a worker be "available" for work in some sense, although some states, such as South Carolina and West Virginia require only that a worker be available to work in his or her usual occupation. Nebraska, New Jersey and Wisconsin, in contrast, require that a worker be available for any work. (Several states require that a worker be available for "suitable" work, although none of these was in the DCA Pilot Project.)

The other main feature of nonseparation eligibility criteria that is shown in Table 2-2 is the handling of workers who refuse suitable work. In all of the pilot states, a worker who refuses suitable work is disqualified from receiving benefits for at least some time, and in two of the states – South Carolina and Wisconsin – such a worker is disqualified for the duration of his or her unemployment spell. In addition, Nebraska and West Virginia reduce the benefits that a worker receives at the end of a disqualification period stemming from refusal of suitable work.

2.4. Decisions Guiding the Design and Reasons for Them

The DCA Pilot Project was developed with the following policy decisions in mind:

- The Department of Labor, following the recommendation of the State-Federal group that developed the Ul Performs design, is committed to developing an approach to measuring denials accuracy;

- There is a need to measure the accuracy of denied claims in the same way as paid claims; that is, there is need to measure the extent to which decisions to deny UI benefits are being made in accordance with fully-informed application of State law and policy;

- The recommendation to use the QPI instrument for reviewing nonmonetary denials accuracy is a reasonable one. Accordingly, it makes sense to ascertain whether a QPI review would be a satisfactory alternative to (or substitute for) the more extensive (and more expensive) full BQC investigation;

- Data on denied claims will be verified in the same way as data on paid claims. This means that the same "full flexibility" in using phone, mail, and fax will be allowed investigators determining the accuracy of denied claims as is allowed in verifying paid claims.

These decisions imply that the DCA Pilot Project was an operational pilot, not a feasibility study or benefit-cost study. The only benefit-cost aspect of the pilot concerned the cost-effectiveness tradeoff regarding use of the QPI vs. the BQC methodology in assessing the accuracy of nonmonetary denials.

It follows that the goal of the DCA Pilot Project was to achieve the following objectives:

- Assess the current range of denial errors and establish desirable sample sizes;

- Determine how much the BQC field-verification methodology adds to the QPI methodology in assessing the accuracy of nonmonetary denials;

- Determine the cost of measuring denials by both the BQC and QPI methodologies;

- Identify the principal reasons for denial errors and ensure that these reasons are

coded so that all states can identify and eliminate or mitigate them;

- Identify obstacles to measuring denials errors consistently.

2.5. Design of the DCA Pilot Project

2.5.1. Overview. The pilot involved five participating states – Nebraska, New Jersey, South Carolina, West Virginia, and Wisconsin – selected from volunteers responding to a system-wide announcement. Each pilot state designated a coordinator who assisted the National Office (NO), the Regional Office (RO), and contractor staff (PRAMM Consulting Group, Inc.) in refining the design and outline materials. The National Office, in collaboration with the states, produced the operational instructions resulting from a series of annual meetings, starting in 1995, that brought together all the participants.

Department of Labor staff drew up the specifications for the universes of denials ("transactions files") from which samples were to be drawn. National Office technical support staff designed the sample selection programs and database programs as variations of the programs already written for BQC and RQC applications. Programming staff in two of the pilot states, South Carolina and Wisconsin, wrote the COBOL program used to edit the populations and select the samples. In addition, programming staff in each state installed and maintained the programs. After training the participating State staff (drawn mainly from existing BQC ranks) and completion of programming, states began sampling and investigating denied claims during September 1997, and continued sampling and investigating over the following year.

Federal and contractor staff monitored each state during the project, and a site visit

was made to each state (except for South Carolina) for the purposes of monitoring the progress of the project. After completion of the investigations, the states completed case data collection instruments, which the Department downloaded to a central database in the National Office. The Department provided copies of the data to the contractor for the evaluation.

In November 1997, staff of the pilot states, the National and Regional Office Federal staff, and contractor staff met in Charleston, South Carolina, to discuss operational issues identified in the first months of the pilot. In November 1998, there was a debriefing meeting to discuss preliminary findings of the pilot and to help shape the final report and its recommendations. States will be asked to review and comment on the draft evaluation report.

2.5.2. Implementation of the Design. Implementation of the pilot involved the following steps:

1. Sampling Cases for Investigation. Starting in week 36 of 1997 (early September 1997), each participating state selected weekly samples of monetary, separation, and nonseparation denials for investigation. The goal was to sample four of each type of denial each week for investigation, and this goal was achieved in all five states by the second week of the pilot. The data elements that were included are indicated in Appendix 1 (the Data Collection Instrument).

2. Investigations. The pilot coordinator/supervisor assigned the sampled denials for investigation.

Monetary denials were investigated using a BQC-like protocol, which involved review of all pertinent agency records, an interview with the claimant and contacts with base-period employers to ascertain the correct wages, hours of work, weeks of work, etc., as prescribed by State law.

Nonmonetary denials (both separations and nonseparations) were reviewed as follows: (a) The DCA-trained investigator conducted a BQC-type review of the case, involving a claimant interview and appropriate contacts with employers and/or third parties to determine the eligibility decision that would accord with a fully-informed application of State law and policy. The field-verification data were coded into a data record maintained on the state's Sun computer. (b) A different investigator performed a QPI appraisal of all separation and nonseparation denials that were drawn for the pilot project. This investigator reviewed all pertinent and available agency records and scored the quality of the denial decision using the existing Quality Performance Indicator (QPI) review instrument. The QPI scoring elements were coded into a data record maintained on the state's Sun computer.

3. Collection of Time Data. Investigators and supervisors were expected to record time required to complete investigations and review cases. In general, each pilot state assigned a project code for the DCA pilot, and State staff recorded time spent performing DCA investigations using the state's time-charge system. Time-recording was kept to an absolute minimum.

4. Case Reviews. As now required under the BQC program, the supervisor was

responsible for reviewing all cases before they were considered final. In addition, a sample of cases was reviewed by Federal or PRAMM monitors during site visits.

5. Data Pickup and Integrity. The Department of Labor picked up the data electronically on a periodic basis for its own use and for transmission to the contractor. The contractor also contacted State staff occasionally in the course of using the data.

2.5.3. Duration of the DCA Pilot Project and Level of State Effort. The pilot involved sampling over approximately a one-year period beginning week 36 of 1997 (early September 1997) and ending between week 33 and 36 of 1998 (late August to early September 1998). South Carolina and Wisconsin completed sampling in week 33 of 1998, New Jersey and West Virginia completed sampling in week 35, and Nebraska completed sampling in week 36. It was initially expected that approximately 3 months would be needed after the last sample was drawn to complete verifications, coding, and case review. (In fact, the data on which the results in Chapter 4 are based were current as of early November 1998 and were not yet complete.) The State pilot coordinators were available for refining the design of the pilot, training, and comment from the execution of the pilot agreement (September 1996) through the debriefing meeting in November 1998 and subsequent review of the contractor's draft evaluation report.

At the beginning of the DCA Pilot Project, a decision was made to provide each state with resources for two DCA Pilot Project investigators. Based on the 1987 Denials Quality Control pilot project, DOL and the states estimated that two investigators would be enough to investigate 200 monetary, 200 separation, and 200 nonseparation denials over

a period of a year. Based on this estimate, the participating states agreed to draw and investigate 200 of each type of denial. However, the initial estimate turned out to be optimistic; in fact, the states found that two investigators could handle only about 150 of each type of denial over the course of a year.

Resources for the DCA Pilot Project were provided to states by two means. First, through the new UI Performs Grants Allocation, the Department of Labor provided resources to each participating pilot state for two DCA Pilot Project investigator positions. Each state provided UI-qualified investigators for the project, usually experienced BQC investigators, and covered travel and other costs involved in investigations. Second, resources for the pilot coordinator (including travel) were provided through a cooperative agreement between the Department of Labor and the states. In addition, each state was provided with $15,000 through the pilot agreement to cover programming costs, which included developing the transaction files and installing and maintaining the sampling and database software. South Carolina and Wisconsin received additional funding to cover the cost of writing sample selection programs according to DOL specifications. These programs were shared with the other DCA Pilot Project states and were provided to all states for the purposes of sampling denials and bringing the BQC software into Y2K compliance.

2.6. Timing of the Pilot Project

A kickoff meeting, which was attended by representatives of the five pilot states, the

Regional Offices, the National Office, and the contractor, was held in Washington, DC, in October, 1996. Several decisions were reached at that meeting, including a division of tasks between South Carolina and Wisconsin in producing the COBOL program that would be needed to sample denied claims during the pilot. The Division of Information Technology produced basic database software during summer and early fall of 1997. Sampling for the pilot began in September 1997 and ran through August or September 1998 (depending on the state). A mid-project meeting of all pilot participants, hosted by the South Carolina agency, was held in Charleston, South Carolina, November 12-13, 1997. During the spring and summer of 1998, representatives of the Department of Labor (both National Office and Regional Offices) and the contractor staff visited West Virginia, Nebraska, Wisconsin, and New Jersey. The states' investigations were completed by early November 1998, and a close-out meeting of all pilot participants, hosted by the New Jersey agency, was held in Atlantic City, New Jersey, November 18-20, 1998.

3. Description of the Pilot's Operations and Findings from Monitoring Visits

3.1. Conduct of the Site Visits

During the year of sampling and investigations for the DCA Pilot Project (that is, between September 1997 and August 1998), representatives of the National Office, Regional Offices, and the contractor (PRAMM Consulting Group, Inc.) visited four of the five participating states. These site visits had three main purposes: (a) to inform the staffs of National Office, Regional Offices, and the contractor of the procedures being followed by each state, (b) to allow each participating state to advise the National Office, Regional Offices, and the contractor of problems and concerns, (c) to help ensure the integrity of the DCA Pilot Project and its outcomes, and (d) to ensure the usefulness of the DCA Pilot Project to national implementation of DCA measurement.

No site visit was made to South Carolina, although a Mid-Project Meeting was held in Charleston, South Carolina, on November 13 and 14, 1997, shortly after the start of sampling and investigations for the pilot. The meeting brought together representatives of each participating state, the National Office, the Regional Offices, and the contractor, with the goals of discussing findings, operational issues, and any problems at an early stage of the pilot. As noted in the Mid-Project Meeting Summary (Clark and Skrable 1997), the meeting resulted in agreement on the handling of several operational problems that arose

early in the pilot.

South Carolina was a participant in the 1987 pilot and as a result played a leading role in the 1997-98 DCA Pilot Project. It was clear from discussions with the South Carolina staff at the Mid-Project Meeting in Charleston that sampling and investigations were going smoothly in South Carolina. As a result, South Carolina was given a somewhat lower priority for a site visit than the other four states. The South Carolina staff would clearly have preferred that a site visit be made to Columbia, so that they could express any concerns that they had and provide direct evidence to the National Office, Regional Office, and contractor that their implementation of the pilot was sound. In the end, however, the contractor, in consultation with the National Office staff, decided that an additional visit to South Carolina would not be essential, given the clear evidence that South Carolina's operations were running smoothly.

This chapter reviews the activities of the monitors and State Employment Security Agency staff during each of the site visits and draws some general conclusions from those visits.

3.2. New Jersey Site Visit

On April 21 and 22, 1998, Andy Spisak of the National Office visited the New Jersey Department of Labor to discuss with agency staff their experience in conducting the DCA Pilot Project (Spisak 1998). On the morning of April 21, he met with Charles Salmon (Chief, Benefit Payments), Tom Hynes (BQC Supervisor), Jim Fink (DCA Project Officer), and the

investigative staff – Phyllis Clinebell, Carolyn Merrill, Kelly Sacks, Madeleine Sekulich, and Lucy Webb. The meeting was a free-form discussion of the project management's and investigator's experiences with and observations concerning the DCA Pilot Project. The main points raised in the meeting follow.

1. Staffing and Time Needed for Investigation. New Jersey found that assigning two investigators to DCA was inadequate. A total of five BQC investigators had participated in the investigations of the sampled denied claims. The agency estimated that 9 to 11 hours were required to complete a DCA case, in contrast to about 13 hours for the investigation of a paid claim using the BQC methodology.

Although the use of telephone, fax, and mail to collect data were expected to save time (compared with an in-person investigation), the savings often did not materialize due to call-backs necessitated by non-response, difficulty in securing cooperation over the telephone, and ignored or returned mail.

2. Alternative Base Period Issues. New Jersey law allows UI claimants to qualify for benefits under several alternative base period options. Initial claims are often ruled ineligible based on minimum qualifying wages of $2,020 or 20 weeks worked. Claimants with less than $2,020 or 20 weeks worked are classified as "indeterminate" pending collection of additional base period wage information. Claimants who do not quality under the initial test may qualify when wages from the lag quarter and current quarter are added.

Because of the alternative base period provision, some ineligible claims that entered the DCA monetary samples had been determined eligible by the time the DCA

investigation took place. The implications for national implementation of DCA affect the timing of the sampling for monetary denials.

3. "Pro Forma" Filings. An estimated 15 to 20 percent of the monetary denials in New Jersey resulted from applicants for means-tested welfare who are required to apply for UI, even though they have earned no wages in the base period. In addition, some nonseparation denials result from claimants who wish to establish a disability claim under Social Security or Supplemental Security Income, and who file a UI claim with the expectation that they will be adjudicated unable to work due to disability or illness. (In one case that Spisak reviewed, the claimant actually returned an uncashed check to the local UI office with the statement that she was physically unable to work and was not entitled to UI benefits. She refused to participate in the DCA review.)

4. Training. The BQC and DCA staff agreed that, although the transition from investigating paid claims to investigating denied claims was relatively easy, training for both paid and denied claims would be very helpful, especially regarding nonmonetary (separation and nonseparation) adjudications. Also, staff turnover since the last comprehensive training for BQC investigators reinforces the need for training new investigators.

5. Interstate Claims. The investigators agreed that investigating denied interstate claims was no more difficult than intrastate claims and were in some respects easier, since documentation could be obtained from a central unit that processes interstate claims. A large proportion of the interstate workload for New Jersey consists of claims from Puerto

Rico filed by seasonal agricultural workers.

6. Quality Performance Indicator. The New Jersey staff felt that the results of QPI scoring of the separation and nonseparation denials sampled was largely consistent with the results of the benefit quality scoring of nonmonetary determinations reported on the ETA 9056 quarterly report. The sense of the group was that continuing QPI as part of the national implementation of DCA would be redundant.

7. Case Documentation. Difficulty in obtaining case documentation from the local offices (an issue raised at the Charleston Mid-Project meeting in November, 1997) was eased by having the case investigators pick up agency records at the local offices (rather than requesting records be sent from the local offices). Using several investigators permitted coverage throughout the state.

8. Data Collection Instrument and Coding. The agency recommended that the Error Amount field be expanded from three to four digits to accommodate DCA underpayments that sometimes exceed $999.

In addition to meeting with the BQC and DCA staff, Spisak reviewed approximately 4 to 6 cases of each type of denial, including a mix of decisions determined to be correct and in error by the DCA staff. Spisak concluded that the overall quality of the investigative case work was excellent, particularly with respect to the thorough documentation of the cases. Each case included a concise and clear narrative summary of the case prepared by the investigator. In several cases the claimant could not be contacted, as evidenced by certified letters which were returned as undeliverable, and investigators' notes of

attempted telephone contacts. In some cases, employers could not be contacted, for example, in the case of seasonal employers at the beach resorts which were closed by the time an attempted contact was made.

9. Multiple Issue Cases. Some claims involve multiple issues. Although the issue that was selected for DCA investigation may have been decided in error, the claimant might still be ineligible due to some other issue or issues. For example, in one New Jersey case that was reviewed, the DCA investigation determined that the claimant should not have been denied benefits due to failure to report for an Eligibility Review Interview. However, the investigation also discovered an issue of adequate work search for which the claimant would have been denied benefits.

In the BQC paid claims database, the data element "Amount That Claimant Should Have Been Paid" was established to capture more accurately the net effect of multiple issues. Consideration should be given to establishing a similar element for DCA cases. It should be made clear that DCA investigators are not required to search for collateral issues. However, multiple issues that are discovered in the claimant's file or through the claimant interview should be taken into consideration in coding the case. A data element which reflects the effects of multiple issues may allow managers and analysts to examine the financial implications of an erroneously decided issue.

On the afternoon of April 22, Spisak had a final meeting with Mike Malloy (New Jersey UI Director), Charles Salmon, and Jim Fink. During this meeting, discussion focused on the points summarized above.

3.3. West Virginia Site Visit

On June 11 and 12, 1998, Terry Clark and Julius Greene (of the National Office), Frank Wannamaker (of the Philadelphia Regional Office), and Wayne Vroman, Steve Woodbury, and Woody Wright (all of the contractor) visited the West Virginia Bureau of Employment Programs in Charleston to review the progress of the DCA Pilot Project there (Clark 1998). The group met with Dennis Redden, the project chief for the DCA Pilot Project, and Charlotte Grey of his staff. In addition, several other members of Mr. Redden's staff met with the group at appropriate times during the visit.

Dennis Redden provided a well-structured overview of the eligibility review and denial process as it exists in West Virginia, and reviewed several actual cases to give a sense of content and process. The group concluded that the cases were well done, and the documentation very complete. The group also verified that the QPI review was performed by QPI staff, separately from the DCA investigation. Finally, the group noted that Mr. Redden's staff in Charleston reviews every determination issued by the agency for general accuracy and returns those that are unacceptable to the originator for correction prior to initial release. As such, it is not surprising that the agency has high quality nonmonetary determinations across the board.

The following issues were discussed in some depth during the meetings.

1. Staffing and Time Needed for Investigation. The West Virginia staff expressed the view that insufficient staff time had been allowed for DCA investigation. This was

especially true given that the QPI component was still in place, though it was noted that the QPI component probably would not continue after the pilot is complete.

The West Virginia staff stated that handling a case load of 600 denials per year would require three positions (not two, as were allocated for the DCA Pilot Project). Dennis Redden's findings indicated that the review itself takes roughly 3 hours per case, with data entry requiring between 15 and 30 minutes, and supervisory review another 30 minutes.

2. Response Rates. Difficulty in obtaining responses from claimants was an issue because there is no real incentive for a denied claimant to cooperate in the investigation. West Virginia had found (as expected) that if the claimant had filed an appeal, he or she does respond to the DCA investigation. West Virginia investigators had also found that sometimes a claimant's story would change upon review – since the claimant knew that what was said the first time resulted in benefits being denied. The West Virginia staff call this the "second time around" syndrome.

Even so, West Virginia reported a response rate to the DCA inquiries of over 80 percent. The process involved mailing a questionnaire upon case selection, inviting response, and following up 2 weeks later by telephone if no response had been received. It would appear that this process was successful for West Virginia. The staff felt that is was easier to get a response in a nonseparation investigation than in a monetary or separation investigation because nonseparation denials are temporary. The expectation that additional benefits will be paid in the current benefit year in the case of a nonseparation denial creates an incentive for the claimant to cooperate in the DCA investigation.

3. Time Lapse Standards for Case Completion. The BQC program has set standards for the percentages of cases completed within 60 and 90 days. The group discussed completion targets in view of the different workload burden of DCA. Since DCA is designed to review only a portion of the data that the BQC program requires, it was felt that a shorter target for completion might be appropriate.

4. Timing of Sampling Denials. In the pilot design, it was decided to pull monetary denials immediately after the denial was issued. This created a problem for West Virginia because automatic denials – that is, denials that are issued before Federal or out-of-state wages can be added to a claimant's record – are sampled for DCA investigation. Often, these cases are eligible claims that would correct themselves in a few days when the missing (but already requested wages) are added to the claimant's record.

The West Virginia staff suggested that a slight delay – perhaps 2 weeks – in the selection process would solve this problem. The group agreed that this issue is not unique to West Virginia and needs to be resolved before nationwide implementation.

5. Interstate Claims. The experience of DCA investigators in West Virginia suggested that Interstate Claims could be included in the DCA investigations without difficulty. This was in large part because states are now allowed to use the telephone, mail, and fax to collect data and carry out investigations, rather than relying exclusively on in-person methods.

6. Data Collection Instrument and Coding. The West Virginia staff indicated some coding refinements that need to be made. In particular, some thought needs to be given

to increasing the number of elements in the data base that can be precoded.

7. Study of Informal or "Counter" Denials. Of particular interest was West Virginia's attempt to gather data about the incidence of self denials. The process of claimants selecting themselves out of the system before officially filing a claim for benefits occurs, presumably, in all states, but it is difficult to get information about the process because, by its nature, the process leaves no record to review. As noted in Chapter 1, denials that occur after only an informal inquiry into monetary eligibility (and without a formal claim being filed) are referred to as "informal" or "counter" denials.

In West Virginia, each informal inquiry into the system ("Would I be eligible?") creates a record that can be monitored manually. (The West Virginia agency refers to these as "pseudo-monetary" determinations.) West Virginia volunteered to monitor informal or "counter" claims for nine weeks, and created a program that precluded double entries if more than one inquiry was made. The program also edits out entries for which a formal claim is subsequently (that week) filed. This effort provided what appears to be the first data available on informal denials.

Table 3-1 displays the findings that West Virginia had accumulated after 9 weeks. The data suggest that, during a nine-week period in April and May, 1998, there were 752 informal denials and 551 formal monetary denials of UI benefits in West Virginia. However, none of the informal denials would be selected for DCA review, since only claimants who file a formal claim and are denied benefits become part of the population of denials from which samples are drawn.

While this special study gives an idea of the number of individuals who may be experiencing informal benefit denial, it does not yield data on the accuracy of those informal denials. West Virginia offered to investigate the accuracy of informal denials using the DCA Pilot Program methodology if the Department of Labor would continue to fund the two DCA Pilot Program positions during the year between the end of the Pilot and nationwide implementation of DCA. The group agreed that this would be an excellent opportunity to obtain data that would otherwise be unavailable. However, West Virginia later concluded that significant modification of their data base would be necessary to conduct investigations of informal denials.

8. Appellate Finality. A troublesome issue with the DCA investigative process is that the finality of a state's decision means that DCA investigative staff is unable to correct some instances of denial error. When discovery of an error works against a claimant, as it usually does in a BQC investigation, finality of a decision seems more acceptable, since finality implies that the claimant is not harmed by the outcome of the investigation. (That is, as a result of finality, the erroneous payment stands even though is was discovered to be erroneous.) In DCA, however, West Virginia has found that some erroneous denials can not be reversed because of finality. These are denials that, based on DCA investigation, are erroneous, should be reversed, and that would make the claimant eligible for payment if they were reversed. During the meetings, the group discussed the case of a woman who had been terminated (not for cause) and who had been told by the terminating employer to tell the UI office that she had quit voluntarily because it would look

better on her record. The DCA investigation concluded that the woman should have been eligible for benefits, since the employer admitted giving the incorrect advice (that is, to say she had quit), which in turn had resulted in a denial of benefits. But the denial was final and could not be reversed as a result of the DCA review. The group concluded that the states may need to consider legislative remedies to allow limited jurisdiction when DCA is implemented nationwide.

3.4. Nebraska Site Visit

On June 29 and 30, 1998, Burman Skrable (of the National Office) and Steve Woodbury (of the contractor) visited the Nebraska Department of Labor offices in Lincoln to discuss Nebraska's experience in conducting the DCA Pilot Project. On the morning of June 29, they had a short meeting with Don Gammill, UI program evaluation administrator, and Ron Joyce, head of the BQC Unit, to discuss the general progress of the DCA Pilot Project in Nebraska. They then spent the remainder of the morning with Ron Joyce, who gave them a thorough review of the pilot's implementation in Nebraska. During the afternoon of June 29 and the morning of June 30, Skrable and Woodbury reviewed approximately 20 cases of each type of denial, randomly selected from the DCA files of the BQC unit. They found that the cases were well-documented and easy to follow. Ron Joyce made himself available to answer questions about Nebraska law and practice, and the process was very instructive. Skrable and Woodbury concluded that the quality of the investigations was excellent.

During the site visit, Ron Joyce raised the following issues, most of which he has since elaborated on in his "Summary Report" (Joyce 1998).

1. Staffing and Time Needed for Investigation. The design of the pilot called for two investigators to be assigned and dedicated to the DCA Pilot Project. Originally, the Nebraska staff had expected that separation and nonseparation investigations, since they tend to be "issue-specific," would be completed more quickly than monetary investigations, which tend to be more involved and take longer to complete due to the need to obtain wage records. These expectations turned out to be incorrect: separation and nonseparation investigations took nearly as long as monetary investigations, and a backlog of incomplete cases soon appeared. To help solve the backlog, the DCA case load was distributed among all BQC program staff. This did not solve the basic problem that two DCA investigators were insufficient to handle the case load, but the overload was distributed among five investigators rather than between just two. The clear conclusion in Nebraska, as in other states, is that 600 DCA cases annually are more than two investigators can handle.

The Nebraska staff has suggested that mandating the same DCA sample size for large and small states alike may put a relatively large burden on small states, particularly in regard to meeting timeliness standards (Joyce 1998).

2. Response Rates. Nebraska investigators experienced little difficulty in obtaining information for the investigation from employers. Obtaining information from denied claimants, however, was more difficult. Nebraska staff has estimated that its response rate

from claimants was 74 percent, even with repeated attempts to reach the claimant by telephone.

Nonresponse delayed the completion of investigations because investigators were required to apply the standards established under BQC, which include second notification and subsequent mailing of questionnaires.

3. Timing of Sampling Denials. In the DCA Pilot Project, samples of denied claims were drawn at the end of the most recent production run, meaning that records were selected before claimants had been notified of the reason for their denial. This proved to be a problem in Nebraska because claimants could not be contacted until they had been duly notified. The result was further delays for case completion. Nebraska staff have suggested a procedural change that would solve this problem: Download DCA samples and distribute them for investigation only after the successive sample has been run. This would provide all parties adequate time to receive formal notification, and would also accommodate the potential filing of appeals before being contacted by BQC/DCA staff.

A further sampling issue raised by Nebraska staff concerned monetary denials that involve combined wage claims. Specifically, Nebraska staff suggested that a combined wage claim should not be considered monetarily ineligible until the State Employment Security Agency has formally reached that conclusion. This would avoid drawing cases that, although they appear to be monetary denials at the time they are drawn, turn out to be eligible for benefits.

4. Programming. Nebraska was able to adapt the COBOL II sampling program

without difficulty. A few sampling errors were found early in the pilot, and the problems were corrected. In general, programming and sampling appear to have gone quite smoothly in Nebraska.

5. Data Collection Instrument and Coding. Nebraska staff has noted that the Data Collection Instrument for the DCA Pilot Program was derived from the existing BQC Data Collection Instrument. As a result, coding the findings of investigations was a problem in certain cases and could be simplified in others. One problem, for example, was that the DCI would accept only three digits for a monetary denial error, which means that the potential dollar error could not be reflected accurately.

The main simplification suggested by Nebraska staff is that the DCI should provide the option to stamp any field (in other words, increase the number of defaults). For example, the DCA has a series of earnings-related fields that do not apply to nonseparation denials, and hence could default to zero. Nebraska staff estimate that there are as many as 26 such fields. Another example would be the "before audit" fields, which do not apply to monetary denials, and could also default to zero. A review of the DCI should lead to a determination of which coding elements are necessary.

6. Dollar Implications of Erroneous Denials. Finally, the Nebraska staff has been especially concerned about the difficulties associated with estimating the dollar implications of erroneous denials. As discussed further in Chapter 5, estimating the dollar impact of erroneous denials is inherently difficult and speculative, and Nebraska's law and practice raise additional issues that are likely to make such estimates rather unconvincing

(see Chapter 5 and Joyce 1998).

3.5. Wisconsin Site Visit

On August 3 and 4, 1998, Jack Estes (of the Chicago Regional Office) and Steve Woodbury and Woody Wright (of the contractor) visited the Milwaukee offices of the Unemployment Insurance Division, Wisconsin Department of Workforce Development, to learn about Wisconsin's implementation of the DCA Pilot Project. Most of their time was spent meeting with John Mand and Chet Frederick of the Wisconsin Department. The visitors also met and talked with the DCA investigators in the Milwaukee office. During the morning of August 4, Estes, Woodbury, and Wright reviewed 4 or 5 DCA cases representing each type of denial, and found them to be well-documented. John Mand was available to answer questions about Wisconsin law and practice. The visiting team concluded that the quality of the investigations was excellent.

During the site visit, John Mand and Chet Frederick raised the following issues, many of which are discussed at greater length in Mand's "Denials Closeout Conference Notes" (Mand 1998).

1. Staffing and Time Needed for Investigation. As was true in other states, Wisconsin found that two DCA investigators could not handle a case load of 600 denials (200 denials of each type) per year. Rather, it appeared that 3 investigator positions would need to be allocated to manage such a case load. Wisconsin estimates that 2 investigator positions would be able to support an annual sample of 450 cases (150 of each type of

denial). An allocation for supervision and support staff would also need to be included.

Wisconsin, again like other states, distributed its DCA cases to all of its investigators, rather than having two DCA specialists, as had been envisioned originally. An advantage of using all investigators for DCA was that it provided for cross-training and backup of staff who might be absent.

2. Time Lapse Standards for Case Completion. Wisconsin favored timeliness standards for Denials, similar to the standards established for BQC cases. However, there was a concern that the higher level of nonresponse among claimants in the DCA Pilot Project may create greater problems in meeting time-lapse standards in the denied claims accuracy program than in the existing BQC (for which response rates are higher). For BQC cases, Wisconsin has experienced response rates around 98 percent. In contrast, Wisconsin's claimant response rate during the DCA Pilot Project was about 78 percent. Because of this, Wisconsin "reopened" a higher proportion of DCA cases than of BQC cases, as claimants eventually responded (for example, when they later establish eligibility). Under BQC, reopening a case (for example, to recode the claimant's prior "missing" data elements) resets the timeliness clock, occasionally losing the 60 or 90 day timeliness that had previously been met on the case. With DCA, the higher proportion of cases "reopened" would make it unlikely that BQC timeliness standards could be met. Two possible approaches would be (a) to set more lenient standards, or (b) add a "reopen code" (one that does not reset the timeliness clock) for the purpose of recoding a case due to a claimant's later response.

3. Timing of Sampling Denials. Other states felt that problems could be avoided by sampling denials (particularly monetary denials) one to two weeks later than was done during the DCA Pilot. Wisconsin staff, however, did not share this view. Rather, Wisconsin staff favored sampling denials immediately after each week's production has been completed, with no lag in sampling. This is the existing procedure with BQC, and was the procedure for the DCA Pilot Project. Wisconsin staff suggest that, with no delay in sampling, they were often in contact with the parties to the denial decision within the initial 14-day appeal period. This gave investigators the opportunity to answer questions, explain the decision and its effect, and advise the parties as to their appeal rights should they choose to exercise them.

4. Data Collection Instrument and Coding. Wisconsin staff had a variety of suggestions for improving the DCI. Some of these suggestions concerned the attempt, which is embodied in the DCI, to estimate the dollar impact of erroneous denials. Wisconsin staff agreed with other states on the difficulties inherent in attempting such estimates.

In addition, Wisconsin agreed with other states that many fields in the DCI should be programmed to default to prespecified values for each type of denial investigation. For example, in nonseparation cases, a series of possible data elements should be coded as "not applicable."

5. Programming. Wisconsin reported very little difficulty with programming, and the DCA sample selection program performed well. There was one sampling error during the

first week of sampling, which came from an invalid nonmonetary issue family. The problem was corrected by removing that issue family from the programming. The invalid case was replaced by a new case from the last batch sampled.

6. Quality Performance Indicator. The views of Wisconsin staff on QPI are similar to the views of agency staff in other states. The essence of the view is that QPI merely reviews the information that already exists in a case file and ensures that proper processes were followed in making a determination. A DCA investigation, on the other hand, attempts to verify the facts that are in a case file (which are sometimes found to be incorrect) and to secure missing information and additional information (for example, statements from the claimant or employer that may not be in the file, rebuttals, and other information not previously obtained). The DCA investigation then determines the correctness of a denial in light of the verified and more complete set of information that is then available. The QPI process is viewed as a good indicator of the "quality" of an investigation, but an inadequate indicator of whether a decision is proper.

7. Value of the Investigating the Accuracy of Denials Claims. In his memo summarizing Wisconsin's experience with the DCA Pilot, John Mand states that, "the Denials review has brought to light a number of problems, either in process or procedure, that were not easily recognizable, if at all, from the BQC review. This has much to do with the fact that we're looking at a claim and/or case from different perspectives. A payment case has often overcome some of these problems to get there, while the denials case has found some of these problems to be fatal to ever establishing entitlement/benefit payment.

In concert, I think the BQC and Denials reviews give us a much better overview of the UI program, than either one can individually."

8. Denials Forms. Wisconsin staff noted that, because of the short preparation time in implementing the Denials Pilot, they had adapted their forms and procedures directly from BQC. They believed that these could be refined to be more specific to investigating Denied Claims Accuracy. In particular, they were consolidating their three denial coding "worksheets" into one worksheet, grouping together data elements that are common to all types, those that are specific to one type, and following the input flow of the entry screens and the coding manual.

It appears that most of the Pilot states used three types of Claimant Questionnaires, adapted from BQC, for their monetary, separation and nonseparation denials. Wisconsin's staff has been debating the merits of a single questionnaire adapted from BQC, compared with three separate questionnaires, one for each type of denial investigation.

3.6. Summary of Findings from the Site Visits

This section brings together the observations that were gathered during the site visits and summarizes the states' comments and concerns about the DCA Pilot Project. With only minor exceptions, the states' observations agree with each other. The states' comments were in general made with an eye to nationwide implementation of DCA measurement. Accordingly, this summary attempts to identify the issues that require consideration before nationwide implementation of the DCA program.

1. Staffing and Time Needed for Investigation. The pilot states reported that the time required to complete DCA cases was only somewhat less than BQC paid claims investigations. As a result, they suggest that the samples sizes used in the DCA Pilot Project will need to be reduced for national implementation. Samples of 200 of each type of denial (600 total) were drawn for the DCA Pilot Project, based on an estimated two staff years to complete the cases. The consensus of the pilot states, however, was that 3 staff years would be needed to handle that case load. Sample sizes will have to reflect resource requirements and resource availability.

A related issue is the relative sizes of the DCA and paid claims samples. Although in the pilot, only 200 cases of each denial type were samples, the aggregate DCA samples were 600 cases annually, compared with 480 paid claims in the "core" BQC program. In addition, the ten states with the smallest UI workload currently sample only 360 BQC cases per year. It may be necessary to reconsider DCA sample sizes for the ten smallest states in light of the current BQC sample differential.

2. Response Rates. The pilot states all report lower response rates from claimants in their DCA investigations than they experience in their core BQC investigations. The reason is clearly that denied claimants – especially those denied for monetary and separation reasons – have little incentive to cooperate with the agency. As a result, claimant response rates to the DCA investigations were in the range of 74 to 81 percent (except in South Carolina, where they were 97 percent or higher). This compares with claimant response rates over 95 percent for core BQC investigations. (Note that the

response rates discussed in this section are based on discussions that took place during the site visits. More complete and detailed data on claimant response rates are discussed in section 4.5 below.)

3. Time Lapse Standards for Case Completion. Unlike the core BQC program, there were no standards for timely completion of cases in the DCA Pilot Project. The pilot states reported that difficulty in obtaining responses from claimants was the main cause of delays in completing cases. The pilot states do favor timeliness standards, recognizing that standards for DCA may need to be structured somewhat differently than existing BQC timeliness standards. Wisconsin suggested either setting more lenient standards or adding a "reopen code" (one that does not reset the timeliness clock) for the purpose of recoding a case due to a claimant's later response.

4. Timing of Sampling of Denials. With the exception of Wisconsin, the pilot states agree that the sampling of monetary denials should be delayed for two weeks (ten work days) from the date that the claim was filed. Such a delay would avoid including in the samples to be investigated claims that were initially denied but that will be redetermined based on confirmation of additional wages for Combined Wage Claims (intrastate claims that combine earnings from one or more employers from another state in order to establish monetary eligibility or to yield the highest possible benefit amount for the claimant), interstate claims, and (in some states) alternative base year qualification.

5. Interstate Claims. The pilot states reported no greater difficulties with investigations involving interstate claims than with those involving intrastate claims.

Indeed, in some cases, the pilot states found interstate claims easier to handle because documentation could be obtained from a central unit that processes interstate claims.

6. Quality Performance Indicator. During the site visits, staffs of the pilot states generally indicated that they prefer for intensive investigation of denied claims (as in the DCA Pilot Project) to the Quality Performance Indicator rating the nonmonetarily denials. Their reasons for this preference varied, but in general they believe that the QPI rating, which only scores the "quality" of the process that led to a nonmonetary denial, can be misinterpreted as an indicator of whether denials are correct or incorrect. But the QPI is emphatically not such an indicator. The pilot states, in other words, want to measure the accuracy of denials through intensive investigations, since such a process will result in a measure of the accuracy of denied claims that is what it appears to be.

7. Programming. The pilot states reported only minor (and easily resolved) start-up problems in implementing the COBOL sampling program and getting the DCA Pilot Project running.

8. Data Collection Instrument and Coding. Most of the pilot states had suggestions for changes in the Data Collection Instrument. Many of the suggested changes pertained to allowing fields in the DCI to be programmed to default (or to be "stamped") to prespecified values for each type of denial investigation. Another frequently noted problem was that the Data Collection Instrument accepts only three digits for a monetary denial error. Finally, some states have specific issues that must be accommodated in the Data Collection Instrument. For example, because New Jersey receives many "pro forma"

filings, a separate action code should be established for UI claims filed as a requirement for establishing a welfare, disability, Supplemental Security Income, or other social program claim. This would allow a state and the Department of Labor to identify and remove these cases from the rest of the sample, so that they do not distort estimates of denial accuracy rates. For the most part, it is not possible to flag such cases in the state's database in order to screen them out of the sampling frame. These cases can be identified through the DCA investigation, however, and the determination can be made that the UI claim was filed solely or primarily for the purpose of establishing eligibility in another program.

9. Training. Although the staffs of the pilot states indicated that the transition from the investigation of BQC paid claims to denied claims was fairly smooth, some also believed that training in the BQC methodology in general and in nonmonetary determinations in particular would be beneficial. Due to staff turnover, there are investigators in some states who have been trained "internally" by other state staff; however, the implementation of DCA might be a good opportunity to offer BQC training to those staff who have come into the program since training was last offered and as a "refresher" to staff who feel they might benefit.

4. Quantitative Analysis of Denials Accuracy

This chapter presents basic quantitative results from the Denied Claims Accuracy Pilot Project. The main questions addressed include the following:

- For each of the three types of denials, in each of the five pilot states, what was the incidence of erroneous denials during September 1997 through September 1998? Where are the problems concentrated? How serious is the problem of erroneous denials? These questions are addressed in section 4.1.

- What are the causes of errors? At what point in the investigation process were errors detected? These questions are addressed in section 4.2.

- To what extent are errors corrected by the states' own internal processes? To what extent did appeals procedures correct erroneous denials? These questions are addressed in section 4.3.

- Are there differences in denied claim errors between combined wage claims (that is, intrastate claims for which monetary eligibility depends on wages from at least one employer in another state) and other claims? This question is addressed in section 4.4.

- To what extent does difficulty in obtaining information from claimants affect the estimates of denied claim errors? How great a barrier does claimant nonresponse pose for the measurement of denied claims accuracy? These questions are addressed in section 4.5.

- For nonmonetary determinations, how do the results of the DCA pilot compare with the QPI assessment of the quality of the determinations process? What factors might explain differences between the accuracy of denied claims as measured in the DCA pilot and the QPI assessment? These questions are addressed in section 4.6.

- How do the characteristics of claimants who are erroneously denied benefits compare with the characteristics of claimants who are properly denied? That is, to what extent does the likelihood of erroneous denial vary with claimant characteristics? Is there information in such a comparison that can be used to

help reduce determination errors? These questions are addressed in section 4.7.

One question that is of interest cannot be addressed in the quantitative analysis of this chapter. This is whether there are differences in denied claim errors between interstate claims and intrastate claims. The question is of interest because, although the existing core BQC program is restricted to intrastate claims, the DCA Pilot Project encompassed both intrastate and interstate claims. However, we are unable to address the question because the Data Collection Instrument that was used in the DCA Pilot Project did not provide a code for an interstate claim (it did provide a code for a combined wage claim). Because the pilot states reported that investigating interstate denials is not substantially more difficult that investigating intrastate denials (see Chapter 3), the DCA program to be implemented in 2000 can be expected to include interstate claims. Also, the Data Collection Instrument will be modified to identify interstate claims, so it will be possible to calculate error rates for interstate claims when the DCA program becomes a nationwide program.

[Note on sample sizes: The pilot design called for each state to randomly sample 200 monetary, 200 separation, and 200 nonseparation denials for investigation. The samples analyzed in this section reflect the pilot design except in one case. As noted in section 3.4, in West Virginia, automatic monetary denials were issued – and sampled for investigation – before Federal or out-of-state wages could be added to a claimant's record. Often, these automatic monetary denials were eligible claims that corrected themselves when the missing wages were added to the claimant's record. Since these cases were in

fact eligible monetary claims rather than true denials, they were given a special code and have been deleted from the samples used in the analyses presented in this section. Throughout this section, then, roughly 200 monetary, separation, and nonseparation denials are analyzed for each state, with the exception of monetary denials in West Virginia, where there are 126 cases.]

4.1. Incidence of Erroneous Denials

Table 4-1 displays basic data on case error rates in each of the five pilot states for each of the three types of denials. The bottom row of the table aggregates the experience of all five pilot states. As discussed in chapter 2 (section 2.3), there are significant differences in the UI laws, policies, and procedures of the five pilot states. As a result, summing the denial error experiences of different states is like adding apples and oranges, and comparing the denial experience of one state with another (or with the average of the five pilot states) may not be meaningful. As long as one keeps in mind the differences among the states, however, the average experience of the five pilot states offers a useful benchmark.

There is much variation in error rates by type of denial. Overall, the pilot states report that separation denials are the least likely to be in error, with an average error rate of 8.7 percent. Monetary and nonseparation denials, on the other hand, are more likely to be in error, with overall error rates of 16 percent and 15 percent. This pattern of separation denials being more accurate that monetary and nonseparation denials is quite pronounced

in Nebraska, South Carolina, and West Virginia, but is less clear in New Jersey and Wisconsin.

For all three types of denials, there is considerable variation in error rates across the five pilot states. Monetary denial error rates range from a low of about 10 percent in Nebraska to a high of 23.4 percent in South Carolina. Separation denial error rates range from a low of 4 percent in Nebraska to a high of 19.7 percent in Wisconsin. Nonseparation denial error rates range from a low of 6.8 percent in West Virginia to a high of 21.7 percent in Wisconsin. Many observers believe that these interstate variations are a reflection more of interstate differences in laws and policies than of differences in the ability or efforts of state agencies to apply their law and policies. In other words, laws and policies that are more difficult to apply are more likely to give rise to erroneous determinations about eligibility. Accordingly, ranking the states by their case error rates does not necessarily imply any judgment about the quality of an agency's enforcement efforts. It will be useful to pin down the aspects of the pilot states' laws and policies that are most likely to result in higher case error rates.

4.1.1. Monetary denials. Table 4-2 breaks down monetary denials by reason and accuracy. Possible reasons for monetary denial included in the DCA data collection instrument include insufficient base period wages, insufficient hours, weeks, or days, failure to satisfy the high-quarter wage test, failure to meet a transitional work requirement, and other reasons. Overall, failure to meet the base period wage test is the most common reason for a monetary denial, followed by failure to meet the high-quarter wage test, and

failure to meet a transitional work requirement.

In Nebraska and New Jersey, virtually all monetary denials stem from insufficient base period wages or weeks of work. However, in South Carolina and Wisconsin, a significant percentage of monetary denials stem from failure to meet the high-quarter wage test, and in West Virginia and Wisconsin, a significant percentage of monetary denials stem from failure to meet a transitional work requirement. In South Carolina, monetary denial error rates appear to be similar across the different reasons for monetary denial. In West Virginia, however, there were no errors on monetary denials resulting from failure to meet a transitional work requirement, whereas the error rate on monetary denials resulting from failure to meet the base period wage requirement was about 18 percent. By contrast, in Wisconsin, the error rate on monetary denials resulting from failure to meet a transitional work requirement was about 46 percent, whereas the error rate on monetary denials resulting from failure to meet the base period wage requirement was about 16 percent.

4.1.2. Separation denials. Table 4-3 breaks down separation denials by reason and accuracy. Possible reasons for a separation denial included in the DCA data collection instrument include lack of work, voluntary quit, discharge, and not separated (still attached to a job, on a leave or absence, or the like). Overall, 60 percent of all separation denials occur because of a voluntary quit, and about 39 percent occur because of a discharge.

This pattern – voluntary quits accounting for about 60 percent of separation denials, and discharges accounting for about 40 percent – is consistent across all pilot states

except South Carolina, where the pattern is reversed. Nevertheless, the case error rates for separation denials stemming from discharges and quits are similar within each state. For example, in South Carolina, about 6 percent of separation denials due to a quit were erroneous, and about 4 percent of separation denials due to a discharge were erroneous.

4.1.3. Nonseparation denials. Table 4-4 breaks down nonseparation denials by reason and accuracy. Possible reasons for nonseparation denial included in the DCA data collection instrument include inability to work, unavailability for work, failure to meet the work search test, disqualifying or unreported income, refusal of suitable work, a reporting or registration violation, and other. In the five pilot states combined, reporting and registration violations were the most common reason for a nonseparation denial (about 30 percent), followed by unavailability for work (about 24 percent), disqualifying income (18 percent), and inability to work (12 percent).

There is considerable variation across the pilot states in the importance of these main reasons for nonseparation denials. Reporting violations and disqualifying income dominate in Nebraska. Reporting violations and unavailability dominate in both New Jersey and West Virginia. Inability issues, disqualifying income, availability issues, reporting violations, and failure to meet the work search test are all significant in South Carolina. Unavailability, reporting violations, and disqualifying income are all significant in Wisconsin.

Overall, nonseparation denial error rates stemming from disqualifying income are higher than nonseparation denial error rates stemming from any other reason. This overall

finding appears to result mainly from relatively high nonseparation error rates from disqualifying income in Nebraska and South Carolina. In New Jersey, the nonseparation error rate from reporting violations is relatively high, and in Wisconsin, the nonseparation error rate from availability issues is relatively high.

4.2. Responsibility for Erroneous Denials

It is possible to analyze several additional aspects of the erroneous denials discussed above. First, DCA investigators determined and coded the responsibility for each erroneous denial. Second, investigators determined whether an erroneously denied claim was in error solely because the reason for the denial was wrong. That is, should the claim have been denied, but for some reason other than that given? Third, investigators indicated the point in the investigation process at which the error was detected. Each of these issues is discussed in this section.

In evaluating the findings presented in this section, it is important to keep in mind that the sample sizes generated by the DCA pilot are quite small. Recall that the goal of each of the five pilot states was to draw and investigate samples of 200 monetary denials, 200 separation denials, and 200 nonseparation denials. In turn, the investigations found a total of 144 erroneous monetary denials, 87 erroneous separation denials, and 151 erroneous nonseparation denials in the five pilot states combined. Especially when examining responsibility, actions, and error detection point within a state, the number of erroneously denied cases on which to base inference is quite small. Once Denied Claims

Accuracy measurement is implemented as a continuing program, states will be able pool findings over several years and should be able to draw more reliable conclusions.

4.2.1. Responsibility for errors. Table 4-5 displays tabulations of the responsibility for erroneous denials, as determined by the DCA investigation. The first panel shows responsibility for erroneous monetary denials, the second panel shows responsibility for erroneous separation denials, and the third panel shows responsibility for erroneous nonseparation denials. (Note that fewer erroneous denials are shown in Table 4-5 than in Table 4-1. This is because error issue data are missing for a few of the erroneously denials. For the same reason, Tables 4-6 through 4-9 show fewer erroneous denial cases than does Table 4-1.)

The right-most column shows that responsibility for 78 percent of the erroneous monetary denials was attributable to the agency or to the employer (either individually or together), without any shared claimant responsibility. Specifically, responsibility for 32.6 percent of the errors was assigned to the agency, responsibility for 38.3 percent of the errors was assigned to the employer, and responsibility for 7.1 percent of the errors was assigned to the employer and the agency jointly. The same finding holds broadly across each of the pilot states individually as well – at least 65 percent of the erroneous monetary denials were attributable to the employer or the agency (either individually or jointly) in all five pilot states. In Nebraska, South Carolina, and West Virginia, the employer alone was the most common source of the error. However, in New Jersey, which has a complicated alternative base period law, the agency was responsible most often, and in Wisconsin, the

agency was responsible slightly more often than the employer. Finally, in Nebraska and West Virginia, the claimant (either alone or jointly with the agency) was responsible for over 20 percent of the monetary denial errors (although the samples in both states are quite small).

The main result is that, in most of the pilot states, employer misreporting or failure to report workers' earnings is the most common cause of erroneous monetary denials, with agency errors a close second.

The second panel of Table 4-5 shows responsibility for separation denial errors. The right-most column shows that the agency was responsible for about two-thirds of the erroneous separation denials in the five pilot states. Because the individual state samples are so small in the case of erroneous separation denials, it is difficult to draw strong inferences about the relative importance of employers and claimants in separation denial errors at the state level. Nevertheless, Nebraska and West Virginia attributed no separation denial errors to claimants.

The third panel of Table 4-5 shows responsibility for nonseparation denial errors. As was the case with separation denial errors, responsibility for about two-thirds of all nonseparation denial errors can be attributed to the agency. In Nebraska, New Jersey, and Wisconsin, claimants appear to play a significant (although still secondary) role in responsibility for nonseparation denial errors.

4.2.2. Correct denial but for the wrong reason. Table 4-6 displays tabulations showing the extent to which erroneous denials were incorrect solely because the reason

for the denial was wrong. These tabulations answer the question, should the claim still have been denied, but for some reason other than that given?

The first panel shows that, for monetary determinations, very few claimants were properly denied but for the wrong reason. Nebraska, New Jersey, and West Virginia had no such cases, South Carolina had only 1, and Wisconsin had just 3.

The second and third panels show that, for separation and nonseparation determinations, a somewhat larger percentage of claimants were properly denied but for the wrong reason. This reflects the greater variety of reasons that might lead to ineligibility under separation and nonseparation eligibility criteria. Still, for separation determinations, only 13 claimants (about 15 percent) were properly denied but for the wrong reason in all five pilot states. For nonseparation determinations, just 21 claimants (about 14 percent) were properly denied but for the wrong reason in all five pilot states. For separation and nonseparation determinations, then, the great majority of erroneous denials would have resulted in claimant eligibility if the error found by the DCA investigation had not occurred.

Surprisingly perhaps, Table 4-6 also shows that, in Nebraska, South Carolina, and Wisconsin, the investigations found a significant numbers of erroneous nonseparation decisions that resulted in overpayments to the claimant.

4.2.3. Error detection point. Table 4-7 shows tabulations of the point in the investigation process where the denial error was first detected. Not surprisingly, the error detection point differs significantly across the three types of denial errors. In addition, though, the error detection point varies significantly across the five pilot states within each

of the three types of denial errors.

Regarding monetary denial errors, in all states except South Carolina, the error was first detected either in verifying wages or the conditions of separation, or in examining wage records. In South Carolina, however, the claimant interview was the most common error detection point in monetary denial errors. This likely reflects the greater emphasis South Carolina was able to place on claimant interviews and South Carolina's exceptionally high claimant response rate (see Table 4-12 and the accompanying discussion in section 4.5).

The second panel of Table 4-7 shows that erroneous separation denials were detected mainly through UI records, with verification of wages and conditions of separation and claimant interviews also playing a role.

The third panel of Table 4-7 shows that nearly two-thirds of erroneous nonseparation denials were detected through UI records. Claimant interviews also played an important role in detecting erroneous nonseparation denials.

It should be noted that assignment of the error detection code depends to some extent on the manner in which the investigation is conducted, which in turn depends on choices made by the investigator. That is, whether an error is detected through a claimant interview, through verification of wages and separation conditions, or through examination of wage records depends on the timing (or order) of contacting the claimant, reviewing wage records records, and so on. The pilot states developed different approaches to conducting an investigation (as will be clear in section 4.5), and different investigators may

have different styles of conducting an investigation. These different approaches to investigating erroneous denials can be expected to result in errors being detected at different points in the investigation.

4.3. To What Extent Are Erroneous Denials Corrected by the System?

Ideally, an erroneous denial would be corrected by processes that are part of the UI system – the agency's own checks and follow-up, employers' actions, and appeals filed by claimants who believe that they should be found eligible for UI benefits. To what extent are errors corrected by the states' own internal processes and by the appeals process? This question is addressed here.

4.3.1. Prior agency action. Table 4-8 shows tabulations of actions that were being taken (or had already been taken) by the agencies at the time of the DCA investigation. The first panel shows that there was much variation across the pilot states in prior agency actions on erroneous monetary denials. First, the DCA investigators found that in three of the pilot states (Nebraska, South Carolina, and West Virginia), the error-causing issue could not have been detected through normal procedures for over half of the erroneous monetary denials. In these cases, official procedures had been followed and forms had been fully completed, but the issue was undetectable. This raises the question whether normal procedures are adequate and whether some improvement that would result in greater accuracy might be feasible.

Second, in New Jersey, South Carolina, and Wisconsin, from 30 to 50 percent of

the erroneous monetary denials were in the process of being corrected by the agency. These are cases in which erroneous monetary denials would be corrected by the ordinary workings of the system. The proportion of cases that were being resolved was particularly high – nearly half – in Wisconsin. Roughly one-third of the erroneous monetary denials in Nebraska and West Virginia were in the process of "self-correcting."

Finally, the DCA investigators found that, in between 15 and 55 percent of erroneous monetary denials, the agency (a) identified the issue but took incorrect action, (b) had adequate documentation to identify the issue but did not do so, or (c) did not follow official procedures.

In sum, the first panel of Table 4-8 suggests that, although existing agency procedures would have resolved about 30 percent of the erroneous monetary denials, nearly 40 percent of the errors could not have been detected under existing procedures. This suggests in turn that an examination and possible modification of existing procedures would be worthwhile in some of the states. It appears that the remaining 30 percent of monetary denial errors result from incorrect action or failure to identify the issue or to follow procedures. The most likely solution here is to allocate more resources to training and to take measures to reduce turnover of agency personnel, since personnel with longer tenure are more experienced and skilled in making decisions.

The second panel of Table 4-8 shows the actions taken by states on erroneous separation denials. There are three main findings. First, DCA investigators found that the error-causing issue could not have been detected through normal procedures for about 20

percent of the erroneous separation denials overall. Second, only 8 percent (7 out of 86) erroneous separation denials were in the process of being corrected by the agency at the time of the DCA investigation. Overall, erroneous separation denials appear less likely to be corrected by agency actions than erroneous monetary denials. Third, incorrect action taken by the agency appears to be the most common action taken by the pilot states in erroneous separation denials.

In sum, the findings in the second panel of Table 4-8 suggests that (a) existing agency procedures would have resolved only 8 percent (overall) of erroneous separation denials and (b) relatively few erroneous separation denials were undetectable (20 percent). It follows that, in the five pilot states together, over 70 percent of the erroneous separation denials involved the agency taking an incorrect action, not identifying an issue, or not following official procedures.

The findings on erroneous nonseparation denials are broadly similar to those on erroneous separation denials. The third panel of Table 4-8 shows that nearly 70 percent of the erroneous nonseparation denials involved the agency taking an incorrect action, not identifying an issue, or not following official procedures. Only 22 percent of the erroneous nonseparation denials were undetectable. As with erroneous separation denials, relatively few erroneous nonseparation denials were in the process of being corrected by the agency (under 9 percent).

The findings in Table 4-8 suggest that the majority of erroneous separation and nonseparation denials result from agency error of some kind. Accordingly, allocating

greater resources for training and reduced turnover of agency personnel would be most likely to reduce the incidence of erroneous nonmonetary (separation and nonseparation) denials. In contrast, changes in agency procedures and improved training of agency personnel are not likely to reduce the incidence of erroneous monetary denials by more than a third. (New Jersey appears to be an exception here, since over half of its erroneous monetary denials were attributed to agency error.) Rather, improved employer reporting of wages would be most likely to reduce erroneous monetary denials.

4.3.2. Prior employer action. Table 4-9 tabulates the actions taken by employers on erroneous denials prior to the DCA investigation. The tabulations show whether, by the time of the DCA investigation, the employer had provided adequate and timely information, had not responded, had not been asked for information, or there was no employer-related issue. These tabulations provide further evidence on the extent to which employers who did not provide adequate and timely information are responsible for erroneous denials.

The first panel of Table 4-9 shows that employers' provision of inadequate or untimely information was most likely to be a problem in the case of erroneous monetary denials, and least likely to be a problem in the case of nonseparation denials. Specifically, for about 52 percent of the erroneous monetary denials, the employer either had not responded or had provided information that was inadequate or late. For about 22 percent of the erroneous separation denials, the employer had not responded or had provided information that was inadequate or late. For only 11 percent of the erroneous nonseparation denials, the employer had not responded or had provided information that

was inadequate or late. (For 73 percent of the erroneous nonseparation denials, there was either no employer issue or no information was requested of the employer.)

The findings displayed in Table 4-9 are consistent with the findings on error responsibility displayed in Table 4-5. Table 4-5 shows that about 52 percent of all erroneous monetary denials were found to be the responsibility of the employer, the employer and agency jointly, or the claimant and employer jointly. Table 4-9's finding that 52 percent of monetary denial errors involved inadequate or untimely employer information is wholly consistent with this finding. Similarly, Table 4-5 suggests that about 14 percent of all erroneous separation denials were the responsibility of the employer, the employer and agency jointly, or the claimant and employer jointly. Table 4-9's finding that 22 percent of monetary denial errors involved inadequate or untimely employer information suggests that inadequate or untimely employer information played a role in many (perhaps most) of the errors erroneous separation denials that were the responsibility of employers. Finally, Table 4-5 suggests that about 9 percent of all erroneous nonseparation denials were the responsibility of the employer, the employer and agency jointly, or the claimant and employer jointly. Table 4-9's finding that 11 percent of monetary denial errors involved inadequate or untimely employer information is consistent with the relatively minor role that employers play in erroneous nonseparation errors.

4.3.3. Appeals and redeterminations. To what extent do appeals and redeterminations correct erroneous UI denials?

The answer depends on whether we consider monetary, separation, or

nonseparation denials. None of the erroneous monetary denials that were investigated in the DCA pilot were reversed on appeal or redetermined. Indeed, only 7 of the 901 monetary denials in the entire sample were appealed, and all of these 7 were proper monetary denials. In other words, there was no scope for appeals and redeterminations to correct erroneous monetary denial because none of them was appealed.

Consider next erroneous separation denials. Only 174 (or 17.3 percent) of the 1,006 separation denials were appealed, and only 23 of these were appeals of erroneous denials. (That is, only about 28 percent of the erroneous separation denials were appealed.) As a result, the scope for the appeals process to reduce the case error rate for separation denials is quite limited. Also, at the time the data were transferred to the Department of Labor, only 39 of the appeals of separation denials had resulted in reversals (and of these, 24 were reversals of proper denials), and 64 were still pending. An additional 3 separation denials had been redetermined (all had been erroneous denials). The upshot is that appeals and redeterminations reduced the number of erroneous separation denials in all five pilot states from 87 to 68 – a rather small reduction of 22 percent.

Finally, consider the erroneous nonseparation denials. Only 71 (or about 7 percent) of the 1,004 nonseparation denials were appealed, and only 13 of these were appeals of erroneous denials. (That is, fewer than 9 percent of the erroneous nonseparation denials were appealed.) So the scope for the appeals process to change the case error rate for nonseparation denials is again limited. Also, at the time the data were transferred to the

Department of Labor, just 11 had resulted in reversals (of these, 6 were reversals of proper denials), and 35 were still pending. An additional 3 nonseparation denials had been redetermined (all had been erroneous denials). As a result, appeals and redeterminations reduced the number of erroneous nonseparation denials in all five pilot states from 151 to 142 – a reduction of only 6 percent.

It is worth noting that, even if the pending appeals of erroneous separation and nonseparation denials resulted in eligibility, few erroneous denials would be corrected. Only 3 of the pending appeals of separation denials were appeals of erroneous denials, and only 2 of the pending appeals of nonseparation denials were appeals of erroneous denials.

4.3.4. Denial error rates adjusted for appeals, redeterminations, and agency actions. Tables 4-10A, 4-10B, and 4-10C pull together the above discussion by recalculating the denial case error rates after adjusting for erroneously denied cases that were reversed on appeal, redetermined by the agency, or were in the process of being resolved by the agency at the time of the DCA investigation.

Consider first the unadjusted and adjusted monetary denial error rates shown in Table 4-10A. Column (1) repeats the unadjusted error rate from Table 4-1. Column (2) shows the monetary denial error rate adjusted for the effects of appeals and redeterminations. Since none of the erroneous monetary denials that were investigated in the DCA pilot was reversed on appeal or redetermined, there is no scope for appeals to influence the case error rates for monetary denials. As a result, the error rates shown

in column (1) of Table 4-10A are identical to those shown in column (2) of the same table.

Agency actions, however, would have corrected some of the monetary denial errors, as noted in section 4.2.1 above. The effect of these agency actions on the monetary denial error rate can be seen in column (3) of Table 4-10A. In the five pilot states taken together, the monetary denial error rate falls from 16 percent to 11.2 percent when agency resolutions are taken into account. The effect of agency actions is most dramatic in Wisconsin, where the monetary error rate falls by nearly 9 percentage points (50 percent). Agency actions also reduce the monetary denial error rate substantially in New Jersey and South Carolina.

Column (4) of Table 4-10A shows the combined effect of appeals, redeterminations, and agency actions on the monetary denial error rate. The figures in column (4) are the same as those in column (3), since appeals and redeterminations had no effect on monetary error rates.

Consider next the separation denial error rates shown in Table 4-10B. Column (1) again repeats the unadjusted error rate from Table 4-1, and column (2) shows the effect of adjusting the separation denial error rate for appeals and redeterminations. As noted in section 4.3.3, only about 28 percent of the erroneous separation denials were appealed, so appeals didn't correct many erroneous separation denials. Overall, the case error rate for separation denials falls from 8.7 percent to 6.8 percent as a result of appeals and redeterminations, a decrease of 22 percent.

Agency actions corrected a negligible number of separation denial errors in the

pilot. As a result, the adjusted error rates shown in column (3) of Table 4-10B are only slightly smaller than those in column (1). In the five pilot states together, the separation denial error rate falls from 8.7 percent to 8.0 percent when agency resolutions are taken into account.

Column (4) of Table 4-10B shows the combined effect of appeals, redeterminations, and agency actions on the monetary denial error rate. Because agency actions corrected few erroneous separation denials, the figures in column (4) are quite similar to those in column (2).

Finally, consider the nonseparation denial error rates shown in Table 4-10C. As in the preceding tables, column (1) shows the unadjusted error rate, and column (2) shows the nonseparation error rate adjusted for the effects of appeals and redeterminations. As noted in section 4.3.3, fewer than 9 percent of the erroneous nonseparation denials were appealed, so appeals could not correct many erroneous nonseparation denials. As a result, the case error rates for nonseparation denials falls by less than a percentage point – from 15.0 percent to 14.1 percent – as a result of appeals and redeterminations.

Similarly, agency actions corrected relatively few nonseparation denial errors in the pilot, so the adjusted error rates shown in column (3) of Table 4-10B are only slightly smaller than those in column (1). In the five pilot states together, the nonseparation denial error rate falls from 15 percent to 13.8 percent when agency resolutions are taken into account.

Column (4) of Table 4-10C shows the combined effect of appeals, redeterminations,

and agency actions on nonseparation denial error rates. Together, appeals, redeterminations, and agency actions reduce the error rate on nonseparation denials from 15 percent to 12.9 percent in the five pilot states.

To summarize, appeals and redeterminations did not reduce the case error rate for monetary denials at all, and they reduced the case error rates for separation and nonseparation denials only marginally (by 1.9 percentage points for the former and by 0.9 percentage point for the later). Appeals and redeterminations did not greatly reduce denial case error rates because relatively few of the the erroneous denials were appealed. Indeed, none of the erroneous monetary denials was appealed, 24 of the 87 erroneous separation denials were appealed, and only 13 of the 151 erroneous nonseparation denials were appealed. In contrast, agency actions did reduce the case error rate for monetary denials substantially (from 16 percent to 11.2 percent in the five pilot states taken together), but they reduced the case error rates for separation and nonseparation denials very little (by 0.7 percentage points for the former and by 1.2 percentage point for the later).

The adjusted case error rates provide a picture of the performance of the UI system in determining eligibility for UI benefits that is both more realistic (in that it reflects the checks that are built into the system) and somewhat brighter than the unadjusted case error rates. Adjusting for the effects of appeals, redeterminations, and agency actions causes the monetary denial error rate to fall from 16 percent to 11.2 percent, the separation denial error rate to fall from 8.7 percent to 6.4 percent, and the nonseparation

denial error rate to fall from 15 percent to 12.9 percent (all error rates are taken over the five pilot states). None of these decreases represents an improvement of more than 30 percent, and the adjusted error rates still suggest a need for improvement in the process of eligibility determination.

4.4. Combined Wage Claims and Erroneous Denials

Combined wage claims are intrastate claims for which monetary eligibility depends on wages from at least one employer in another state. Because these claims involve reports from another state agency, they are more complicated to process and hence potentially more prone to error.

This section presents findings on two aspects of the combined wage claims handled by the DCA pilot states: first, the volume of combined wage claims investigated by the pilot states; and second, the case error rates that were found for monetary, separation, and nonseparation combined wage claims.

4.4.1. Volume of combined wage claims. The first two columns of Table 4-11 tabulate the volume of denials involving combined wage claims that were investigated by the five pilot states. The first panel of the table shows that 4.6 percent (41 out of 901) of the monetary denials handled by the pilot states were combined wage claims. But there is substantial variation among the five states: Over 13 percent of West Virginia's monetary denial case load were combined wage claims, whereas no more than 5.6 percent of the monetary denials case loads in the other states were combined wage claims. (New Jersey

appears to have excluded combined wage claims from its monetary denial case load.)

The second panel of Table 4-11 shows that only 4.5 percent (45 out of 1,006) of the separation denials handled by the pilot states were combined wage claims. There is less variation among the five pilot states in the percentage of separation denials that were combined wage claims: West Virginia still has the highest percentage of combined wage cases (with 8.7 percent), and the other four pilot states range between 2 percent and 4.5 percent.

The third panel of Table 4-11 shows that only 2.7 percent (27 out of 1,004) of the nonseparation denials handled by the pilot states were combined wage claims. The variation among the five states in the percentage of nonseparation denials that involved combined wages could be described as substantial but not dramatic: West Virginia again has the highest percentage (5.3 percent), and the other four pilot states range between 1 percent and 3.6 percent.

An important implication of the findings in the first two columns of Table 4-11 is that the number of denials investigated by each state that involved combined wage claims was quite small. As already noted, only 41 of the 901 monetary denials investigated by all five pilot states were combined wage claims, 45 of the 1,006 separation denials investigated were combined wage claims, and just 27 of the 1,004 nonseparation denials investigated were combined wage claims. It follows that the data available to calculate denial error rates for combined wage claims is quite thin, and that comparisons between combined wage claims and other claims should be interpreted with care.

4.4.2. Case error rates for combined wage claims. The third and fourth columns of Table 4-11 show the denial case errors rates that were found for combined wage claims (CWC) and other claims (Non-CWC) in each of the five pilot states, as well as for the five states combined. As already noted, the denial error rates for combined wage claims are based on very small samples. As a result, comparisons between the denial case errors rates for combined wage claims and other claims are best restricted to the totals for the five pilot states.

What does the evidence suggest about the accuracy of denials that involve combined wage claims? The first panel of Table 4-11 suggests that the error rate on monetary denials is considerably higher for combined wage claims than for other claims. These results are dominated by the experience of West Virginia and Nebraska.

It stands to reason that monetary denials involving combined wages might be more error-prone than other denials. Delays and difficulty in obtaining wage information from out-of-state are likely to increase the probability of error.

In contrast, there is little reason to believe that separation or nonseparation denials involving combined wage claims should be any more prone to error than separation or nonseparation denials that are other claims. The evidence in the second and third panels of Table 4-11 tends to support this expectation. For separation denials, the error rate on claims involving combined wages is 6.7 percent, whereas the error rate on other claims is 8.7 percent in the five pilot states combined. For nonseparation denials, the error rate on claims involving combined wages is 7.4 percent, whereas the error rate on other claims

is 15.3 percent in the five pilot states combined.

The conclusion from Table 4-11 is that there is some reason to believe that monetary denials of combined wage claims are more prone to error than are monetary denials of other claims. In contrast, there is no evidence that separation and nonseparation denials involving combined wage claims are any more prone to error than are separation and nonseparation denials of other claims.

4.5. Claimant Nonresponse and Denied Claims Errors

In Chapter 3, attention was given to the methods used to obtain information in DCA investigations and to the problems that investigators encountered in getting responses, especially from claimants who had been denied benefits. However, the evidence discussed in Chapter 3 was anecdotal. The pilot states kept track of and coded both the methods that were used to obtain information in each investigation and whether there was a response. This section examines those data. The first subsection examines the methods used to obtain information and quantifies the extent of nonresponse. The second subsection examines the relationship between nonresponse and whether a denial was found to be proper or erroneous.

4.5.1. Methods used to obtain information and nonresponse. Table 4-12 displays tabulations of the methods used to obtain information from claimants in DCA investigations and whether the claimant responded. The first panel shows methods and

response rates for monetary investigations, the second panel shows methods and response rates for separation investigations, and the third panel shows methods and response rates for nonseparation investigations.

The main methods used to gather information from claimants were in-person interviews, telephone and fax, and the mail. (The Data Collection Instrument is ambiguous as to how fax is coded.) Most of the variation in methods used to obtain claimant information is across state, rather than across type of investigation within a state. In other words, the five pilot states used somewhat different mixes of methods to obtain claimant information, but once a state had selected a mix of methods to obtain claimant information, it tended to use that same mix to obtain information in all three types of investigation – monetary, separation, and nonseparation.

For example, South Carolina obtained information almost exclusively in person and by telephone and had a very low nonresponse rate. This was true for all three types of investigation in South Carolina. (As noted in Chapter 3, South Carolina's extremely low nonresponse rate results in part from its ability to threaten use of subpoena power, which most state BQC staffs do not have.) Wisconsin also relied heavily on in-person interviews, although its use of the telephone was somewhat lower and its nonresponse rate higher.

The other three pilot states relied very little on in-person interviews. Nebraska relied heavily on telephone interviews, reflecting the size and low population density of the state. West Virginia relied heavily on the mail and telephone follow-up, and achieved the second-lowest nonresponse rate in monetary investigations. New Jersey used a mix of in-

person interviews, telephone interviews, and mail to obtain information.

The nonresponse rates vary dramatically across the five pilot states. As already noted, South Carolina had extremely low nonresponse rates (between 1.5 percent and 3 percent) in all three types of investigation. At the other extreme, New Jersey was unable to obtain claimant information for 57 percent of its monetary investigations.

There was somewhat more consistency in claimant nonresponse rates for separation and nonseparation investigations. Leaving out South Carolina, nonresponse rates for separation investigations were in the range of 21 to 26 percent, and nonresponse rates for nonseparation investigations were in the range of 17 to 31 percent, with New Jersey reporting the highest rate of nonresponse.

During the site visits, the states reported that it was easier to obtain claimant information for nonseparation investigations than for monetary or separation investigations. The logic here is that nonseparation denials are usually temporary and short-term. The claimant generally expects to obtain future benefits and hence has an incentive to cooperate with the UI agency. The same is not true for monetary or separation denials. In these latter, the claimant must find reemployment before he or she will be able to establish a valid UI claim. As a result, there is less incentive for a claimant denied for monetary or separation reasons to cooperate with the agency.

Interestingly, though, the pattern of nonresponse shown in Table 4-12 is not wholly consistent with the above logic or the verbal reports of the states. In fact, nonresponse rates were quite similar across all three types of investigation in South Carolina, West

Virginia, and Wisconsin. In Nebraska and New Jersey, the nonresponse rates for monetary investigations were higher than for nonmonetary investigations; however, the nonresponse rates for separation investigations were about the same as (in Nebraska) or somewhat lower than (in New Jersey) nonresponse rates for nonseparation investigations. Further investigation of the correlates of nonresponse could shed additional light on these issues.

4.5.2. Nonresponse and case error rates. Table 4-13 displays tabulations of case error rates divided into two groups: (a) denials for which claimant information was obtained and (b) denials for which there was no response from the claimant.

The data in Table 4-13 show that error rates tend to be lower when there is no claimant response. This finding can be interpreted in more that one way. One possibility is that, when the claimant fails to respond to the agency's inquiry, investigators tend to find that the denial is proper. This may be the correct conclusion, or the only conclusion possible under the circumstances, but it is nevertheless a conclusion that is based on incomplete information. This is another case where further investigation, perhaps a regression analysis like that developed in section 4.7 below, could provide additional useful information.

4.6. Comparing the QPI Assessment with DCA Measurement

One of the main objectives of the Denied Claims Accuracy Pilot was to compare the results of a comprehensive field investigation (that is, DCA) with the Quality Performance Indicator (QPI) assessment of nonmonetary determinations. As discussed in greater detail

in section 1.4.1, the QPI assessment is designed to appraise the quality of the determinations process, whereas the DCA is designed to determine whether the outcome of the process was correct. The QPI focuses on process, whereas the DCA focuses on outcome.

This section first examines how highly the QPI score is correlated with the results of the DCA investigation. It then goes on to ask whether a denial that fails the QPI assessment is more or less likely to be appealed than is an erroneous denial. Finally, the section offers an attempt to understand which component (or components) seem to be causing the QPI assessment to be at odds with the DCA investigation.

4.6.1. Basic results. Table 4-14 shows cross-tabulations of the accuracy of separation denials by whether the denial determination passed or failed QPI. (If the QPI score assigned to a case is 81 or greater, it is considered passing. If it is 80 or less, it fails QPI.)

To understand Table 4-14, consider first Panel A for all five states. Of the 917 separation denial cases that were both investigated by DCA and had QPI appraisals, 603 were determined proper denials by DCA and passed QPI. Also, 54 denial cases were determined improper denials by DCA and failed QPI. But 233 cases that were proper denials (as determined by DCA) failed QPI, and 27 cases that were improper denials (as determined by DCA) passed QPI.

Another way to view the table is look across the the "improper" row in Table 4-14, Panel A for all five pilot states. In order for QPI to be an accurate estimator of the quality

of a separation determination, we would like all (or nearly all) of the QPI scores in the "improper" row to be failures. But, in fact, only two-thirds of the QPI scores for these cases indicate failure.

Consider next the "proper" row in the in Table 4-14 Panel A for all five pilot states. Again, for QPI to be an accurate estimator of the quality of a separation determination, we would like all (or nearly all) of the QPI scores in the "proper" row to be passes. But, in fact, only 72 percent of the QPI scores for these cases are passes.

Table 4-15 shows cross-tabulations of the accuracy of nonseparation denials by whether the denial determination passed or failed QPI. The findings shown in Table 4-15 are similar to those in Table 4-14. Of the 911 nonseparation denial cases that were both investigated by DCA and had QPI appraisals, 607 were determined proper denials by DCA and passed QPI. Also, 78 nonseparation denial cases were determined improper denials by DCA and failed QPI. But 161 cases that were proper denials (as determined by DCA) failed QPI, and 65 cases that were improper denials (as determined by DCA) passed QPI.

Another finding shown in Tables 4-14 and 4-15 is that, in all five pilot states and for both separation and nonseparation denials, the QPI gives an excessively negative view of the extent to which denials are erroneous. In South Carolina, for example, 29 out of 200 separation denials failed QPI, but DCA assessment showed that only 10 of the 200 separation denials were in fact improper. Also in South Carolina, 51 of the 200 nonseparation denials failed QPI, but DCA showed that only 37 of the 200 separation denials were in fact improper. Similar findings apply to the other pilot states.

The results in Tables 4-14 and 4-15 suggest that the correlation between QPI and DCA is highly imperfect. Statistical tests show this to be the case: in only one case is the coefficient of correlation between the DCA investigation and the QPI appraisal greater than 0.3. In all five pilot states and for both separation and nonseparation denials, a significant percentage of denials that were determined improper by DCA passed QPI, and a significant percentage of denials that were determined proper by DCA failed QPI.

To the extent that the correlation between QPI and DCA is positive (although low), QPI is not a wholly misleading predictor of the accuracy of nonmonetary determinations. But QPI is a very noisy predictor of the accuracy of denials. Many denials that pass QPI are improper, many denials that fail QPI are proper, and overall, QPI suggests that the determinations process is less accurate than comprehensive field audits show.

4.6.2. Failed QPI, erroneous denial, and likelihood of appeal. The finding that QPI is poorly correlated with the DCA determination may or may not be fatal to the efficacy of the QPI. It could be argued, for example, that the process of eligibility determination is what matters most, and that a decision that passes QPI is a good decision. Whether a determination is determined correct by an intensive field investigation would not matter, in this view.

In this section, the above argument is put to a simple test. If a decision that passes QPI is a good or fair decision, and a decision that fails QPI is a bad or unfair decision, then claimants should perceive this, and it should be reflected in their behavior. In particular, it should be the case that denials that pass QPI will tend not to be appealed and that

denials that fail QPI will tend to be appealed. Alternatively, if the results of the DCA investigation are more highly correlated with claimants' perceptions of the fairness of a denial, then denials that are found proper by the DCA investigation will tend not to be appealed and denials that are found erroneous by the DCA investigation will tend to be appealed.

Table 4-16 tabulates appeals of separation determinations by whether the DCA investigation found them proper or erroneous, and by whether they passed or failed QPI. The top panel of Table 4-16 shows that erroneous separation denials are somewhat more likely to be appealed than are proper separation denials. This is consistent with the idea that the outcome of the DCA investigation accords with claimants' perceptions of the fairness of separation determinations. The bottom panel of Table 4-16 shows similar, but less strong, results for QPI: Separation denials that failed QPI are somewhat more likely to be appealed than were separation denials that passed QPI.

Table 4-17 tabulates appeals of nonseparation determinations by whether the DCA investigation found them proper or erroneous, and by whether they passed or failed QPI. The top panel of Table 4-17 shows that erroneous nonseparation denials are slightly more likely to be appealed than are proper separation denials. The bottom panel of Table 4-17 shows that nonseparation denials that failed QPI are also slightly more likely to be appealed than are separation denials that passed QPI.

Overall, the results suggest that both the QPI assessment and the outcome of the DCA investigation are consistent with claimants' perception of the fairness of a denial. The

evidence tends to favor the DCA investigation in the case of separation denials and is a virtual dead heat in the case of nonseparation denials. Based on these results, it would be difficult to make the sort of case put forward at the beginning of this section – that the QPI is a superior indicator because it provides a better measure of the fairness of the determinations process. In fact, the outcome of the DCA investigation seems to be a better indicator of claimants' perception of the fairness of separation determinations, and there is little difference between the QPI assessment and the outcome of the DCA investigation with respect to claimants' perceptions of the fairness of nonseparation denials.

4.6.3. Which QPI component is causing the problem? Section 4.6.1 showed that the QPI indicator is only weakly correlated with the outcome of the DCA assessment. It is natural to ask why. The QPI is made up of six components, each of which yields a score. The scores on each of the six components are summed to yield a total score for a decision. The question posed in this section is whether one or more of the components of the QPI is negatively correlated with the outcome of the DCA field investigation, and hence reducing the correlation between the QPI score and the outcome of the DCA investigation.

A simple way to address this question is to regress the outcome of the DCA investigation of a denial (1 if the denial was found to be proper, 0 if it was found erroneous) on the six components of the QPI. Although more than one way of coding the components of the QPI can be imagined, a simple approach is to code each component as 1 if the aspect of the decision being scored was satisfactory (or not applicable) and 0 if it was unsatisfactory. So, for example, if claimant information was scored as adequate

in a case, the claimant information component of the QPI would be coded 1 for that case; if claimant information was scored as inadequate, the claimant information component of the QPI would be coded 0 for that case.

Table 4-18 shows the results of estimating two equations in which the outcome of the DCA investigation (1 if proper, 0 if erroneous) is regressed on the six components of the QPI (with each component coded 1 if it was scored adequate or not applicable, 0 otherwise). The first equation uses the 902 separation decisions for which QPI scores are available, and the second equation uses the 895 nonseparation decisions for which QPI scores are available.

The results for separation denials (in the first column) suggest that two components of the QPI – the adequacy of other information and whether the provisions of the state's law and policy were met by the decision – are strongly positively correlated with the outcome of the DCA investigation. But two other components of the QPI – the adequacy of claimant information and whether the opportunity for rebuttal was provided – are negatively correlated with the outcome of the DCA investigation. In other words, QPI reviewers have a tendency to score claimant information as adequate and to score the opportunity for rebuttal as having been provided in cases where the DCA investigation finds the decision to be erroneous, and vice versa.

The results for nonseparation denials as shown in the second column, and are similar to the results for separation decisions. These later results suggest that three components of the QPI – the adequacy of other information, whether the provisions of the

state's law and policy were met by the decision, and the adequacy of the written determination – are positively correlated with the outcome of the DCA investigation. But the adequacy of claimant information is again negatively correlated with the outcome of the DCA investigation. So for nonseparation decision as for separation decision, QPI reviewers tend to score claimant information as adequate in cases where the DCA investigation finds the decision to be erroneous, and vice versa.

This exercise suggests the importance of obtaining adequate information from the claimant in reaching proper separation and nonseparation decisions. In particular, it suggests the (somewhat disturbing) conclusion that QPI scoring systematically views claimant information as adequate when, arguably, it is inadequate. It seems that there is no substitute for good claimant information in reaching proper decisions, and that an intensive field investigation uncovers information that would otherwise remain hidden.

4.7. Determinants of Erroneous Denials

Do erroneously denied claims differ systematically from claims that are correct? Or are erroneous denials random events that have no systematic component? To the extent that erroneous denials do have a systematic component, it should be possible to direct or "target" resources toward the types of denied claims that tend to be in error and to avert some of the errors. Doing so would be an efficient way of reducing the overall error rate. On the other hand, if erroneous denials occur at random, then the error rate could be reduced only by devoting greater resources overall to the processes of eligibility

determination. Targeting specific types of claims or denials as most likely to be in error would not be possible. Reducing the error rate would entail greater efforts to obtain information from all parties, greater resources devoted to training decision-makers, and more time spent in making each decision.

This section examines the extent to which there is a systematic component to errors that occur in each of the three types of eligibility determinations – monetary, separation, and nonseparation. The approach is to estimate a regression model for each type of determination in which a dummy variable indicating whether a denial was correct (1 if the denial was correct, 0 if the denial was erroneous) is regressed on a group of observable characteristics of each denied claimant:

- age of the denied claimant in years
- gender of the denied claimant (0 if female; 1 if male)
- whether the claimant was a U.S. citizen (0 if not a citizen; 1 if a citizen)
- ethnicity of the claimant (dummy variables for black, Hispanic, Asian/Pacific Islander, American Indian, and nonhispanic white)
- level of schooling completed [dummy variables for less than high school, high school graduate or GED, some college but no degree, college degree (including associates or higher degree)]
- normal wage in the usual job, in dollars per hour (dummy variables for usual wage rate less than $6 an hour, between $6 and $7 an hour, between $7 and $12 an hour, and over $12 an hour).

In addition, to the above characteristics of the claimant, each regression includes a set of dummy variables for state in which the individual claimed UI benefits. (Because of missing variables in Nebraska, the denials are drawn only from New Jersey, South Carolina, West Virginia, and Wisconsin. A dummy variable is included in the regression for each state.) Also, each of the three equations estimated includes a set of variables

specific to type of denial under consideration. In the equation for monetary denials, a dummy variable is included indicating whether the denial involved a combined wage claim. In the equation for separation denials, dummy variables are included indicating whether the denial involved a voluntary quit, a discharge, or some other separation issue. And in the equation for nonseparation denials, dummy variables are included indicating whether the denial involved an able issue, an available issue, a work search issue, a refusal of work, a reporting or registration violation, or some other nonseparation issue. The reason for including these variables in the estimated equation is to see whether denials involving one or another specific issue are more likely to be erroneous.

The results of estimating such "erroneous denial" equations are displayed in Table 4-19. The estimates shown were obtained by Ordinary Least Squares (OLS) estimation using data on denied claims from New Jersey, South Carolina, West Virginia, and Wisconsin for which information on all included variables was complete. Because the dependent variable in this model is a zero-one (binary) variable, a maximum likelihood procedure such as probit or logit is indicated. When the models are estimated by OLS and probit, however, it turns out that the results are essentially similar. Because the OLS results are simpler to interpret, the OLS estimates are reported here.

4.7.1. Monetary denials. The first column of Table 4-19 shows estimates of the erroneous monetary denial equation. Apart from the constant and the coefficient of one of the state dummy variables, only four estimated coefficients in the model are statistically different from zero at the 5-percent level or better: the age coefficient, two of the usual

wage coefficients, and the combined wage claim coefficient. In general, the interpretation of the coefficients in the model is as follows: A unit increase in an independent variable changes the probability that the monetary denial was correct by the amount given by the estimated coefficient. For example, consider the coefficient for usual wage less than $6 per hour: For a worker whose usual hourly wage rate is less than $6 per hour, the probability that a monetary denial is correct is higher by .166 than for a worker whose usual hourly wage is greater than $12 per hour (the reference group). Consider next the age coefficient: For a worker aged 30, the probability that a denial is correct is lower (the coefficient is negative) by .003 than for a worker aged 29. (It follows that the monetary denial of a 40-year-old worker is less likely to be correct by .03 than is the monetary denial of a 30-year-old.) Finally, consider the combined wage claim coefficient: For a combined wage claim, the probability that a monetary denial is correct is lower by .437 than for a denial that does not involve a combined wage claim.

These results suggest that there is at least some systematic component to the incidence of erroneous monetary denials – that erroneous monetary denials are not wholly random accidents. Monetary denials of older workers, workers with a high usual wage, and combined wage claims are all more likely to be in error than are other denials. The result for combined wage claims is not surprising, but the result for age and usual wage may be. The latter results make sense, however, considering that relatively few high-wage and older workers are denied. The results suggest that, conditional on being denied on monetary grounds, older workers and higher-wage workers are more likely to experience

an erroneous monetary denial.

4.7.2. Separation denials. The middle column of Table 4-19 shows estimates of the erroneous separation denial equation. These estimates suggest two points. First, the very few separation denials that involve issues other than voluntary quits or discharges are less likely to be proper by about 0.4 than are the more common separation denials (that is, those that do involve voluntary quits and discharges). Second, separation denials of American Indians are less likely to be proper by about 0.4 than are the separation denials of other racial or ethnic groups. This would be a troubling result except that it is based on just one case. Otherwise, the model suggests that separation denials are random accidents that have no systematic component.

4.7.3. Nonseparation denials. The right-most column of Table 4-19 shows estimates of the erroneous nonseparation denial equation. These estimates yield a conclusion that is similar to the one just drawn of separation denials: Nonseparation denials are largely random events that have no systematic component. There is only one exception: Nonseparation denials that involve disqualifying income are less likely to be proper by about 0.13 than are nonseparation denials that involve other issues.

5. Is It Possible to Estimate the Dollar Impact of Erroneous Denials?

5.1. Description of the Problem

The existing Benefits Accuracy Measurement (BQC) program produces estimates of the overpayments made by the UI system – that is, payments in excess of those that should have been made. Conceptually, at least, it is a straightforward procedure: A UI payment was made; the BQC investigation determines that it should not have been made; the payment that should not have been made is weighted and summed with other overpayments to yield an estimate of the overpayments made by each state and the entire system.

Could the Denied Claims Accuracy program produce parallel estimates? Some consideration suggests that producing estimates of the dollar implications of erroneously denied claims is essentially different. When a claim is erroneously denied, no payment is made so no payment can be observed. For example, in the case of an erroneous monetary denial, would the wrongly denied claimant be fully eligible for UI benefits? (That is, would he or she satisfy the separation and nonseparation criteria for benefits, in addition to the monetary criteria?) How many weeks of benefits would the claimant draw? Would the benefits drawn each week be partial or full? None of these questions can be answered through direct observation. We cannot know with certainty whether the claimant would have been fully eligible, would have started drawing benefits if he or she had been found

eligible, or the number of weeks of benefits that would have been drawn, if the erroneous denial had not occurred.

To summarize, the basic difficulty in estimating the dollar impact of erroneous denials is that the experience of a worker who is erroneously denied benefits is not observed.

5.2. Possible Solutions

This section briefly explores a possible solution to the problem outlined in the previous section. Although the experience of a worker who is erroneously denied benefits is never observed, it may be possible to impute or simulate the benefits that the erroneously denied claimant would have received had he or she not been erroneously denied.

A simple approach to imputing the benefits that would have been received by an erroneously denied claimant would proceed by (a) obtaining data on a sample of eligible UI claimants, observed at the end of their benefit year, and (b) estimating a statistical model in which the dollar benefits received by a claimant during the benefit year depend on the observable characteristics of the claimant (such as work experience, industry and occupation, age, gender, race, and any other relevant and observable characteristics). This estimated statistical model could be viewed as the "structure" by which the dollar amount of benefits received by a claimant is determined. Accordingly, it could be used to impute the dollar amount of benefits that would have been received by a claimant who was

erroneously denied benefits. The procedure would be to substitute into the estimated model the observable characteristics of an erroneously denied claimant. The product, again, would be an estimate (or imputation) of the dollar amount of benefits that would have been received by the erroneously denied claimant.

The above approach is appealing (especially to an economist or a statistician), although it encounters three problems. First, it assumes that the necessary data are available. Second, the approach assumes that all erroneously denied claimants should have been eligible for benefits. Third, statistical modeling assumes that erroneously denied claimants are essentially similar to claimants who are found eligible for benefits – in other words, that erroneous denials are random errors. We consider each of these problems and possible solutions in turn.

5.2.1. Data availability. Estimating statistically the benefits that erroneously denied claimants would have drawn if they had not been denied benefits requires a random sample of eligible UI claimants at the end of their benefit year. Each state's administrative records make it possible, in principle, to generate such data, but the feasibility of actually producing such data on a regular basis has (to our knowledge) not been explored. It is worth noting that the existing BQC data, as currently constituted, could not be used to estimate the sort of model described above because BQC records do not include information of the claimant's experience during the full benefit year. Nevertheless, the BQC sampling frame could be used to obtain the needed data: States could in principle redraw the data on each claimant in the BQC sample at the end of his or her benefit year.

5.2.2. Would erroneously denied claimants have been eligible for benefits?

Second, statistical estimation of the benefits that erroneously denied claimants would have drawn assumes that all erroneously denied claimants would be eligible for benefits if they were not erroneously denied. This assumption could be incorrect for either of two reasons: (a) In some cases, an erroneous denial may be correct but for the wrong reason. For example, the reason given for a separation denial may be wrong, but the claimant would still fail the separation criteria for some other reason. (b) Claimants who are erroneously denied for monetary reasons still need to satisfy the separation and nonseparation criteria for eligibility. And claimants who are erroneously denied for separation reasons still need to satisfy the nonseparation criteria for eligibility. There is no guarantee that either would be the case. The point is that correction of an erroneous decision does not necessarily yield an eligible UI claim. Accordingly, correction of an erroneous decision may have no implications for dollar payments.

Regarding point (a), the evidence presented in section 4.2.2 (and Table 4-6) suggests less than 3 percent of erroneous monetary denials, 15 percent of the erroneous separation denials, and 14 percent of the erroneous separation denials were "proper but for the wrong reason." Accordingly, this problem is more conceptual than practical, although it cannot be ignored

Handling point (b) would probably require a redesign of the Denied Claims Accuracy investigation. As it is currently set up, the DCA program will investigate three random samples of denials – a sample of monetary denials, a sample of separation denials, and

a sample of nonseparation denials. An investigator who is handling a monetary denials will focus on monetary criteria for eligibility, an investigator who is handling a separation denials will focus on separation criteria for eligibility, and so on. Determining whether an erroneously denied claimant would be eligible for benefits would require an all-encompassing investigation of each sampled denial. The feasibility of such an investigation has not been explored, but this issue has prompted several of the participating pilot states to voice their strong concerns about attempting to estimate dollar impacts (see, for example, Joyce 1998; Mand 1998; and Chapter 3 above).

5.2.3. Are erroneously denied claimants essentially similar to eligible claimants? The statistical model that provides estimates of the dollar amount of benefits received by a claimant would need to be estimated using a sample of claimants who were determined eligible for UI benefits. Using that model to impute the benefits that would have been received by erroneously denied claimants requires the assumption that eligible and erroneously denied claimants are essentially similar in all ways, both observable and unobservable. In other words, we need to assume that the erroneously denied claimant was erroneously denied through a random accident, and the error could just as easily have happened to one of the claimants who was determined eligible. Whether this is the case is unknown, although further analysis of the data from the DCA Pilot Project and BQC could shed some light on the extent to which eligible and erroneously denied claimants are similar in observable ways. The limited evidence on this point that is offered in section 4.7 suggests that erroneous denials are random accidents. If further evidence points to the

same conclusion, then modeling of the kind discussed above should be able to proceed.

5.3. Summary and Conclusions

The essential problem inherent in estimating the dollar impact of erroneous denials is that we cannot observe the benefits that would have been received by an erroneously denied claimant had he or she not been erroneously denied. Statistical modeling seems to be an appealing way to overcome the this problem, but statistical modeling encounters three problems. First, the data needed to implement statistical modeling are not currently available. Second, statistical modeling requires the assumption that correct determination of an erroneously denied claim would result in eligibility for benefits. Third, statistical modeling depends on the assumption that eligible and erroneously denied claimants are alike in all observable and unobservable ways.

The first problem, availability of data, can be solved in principle. The question here is whether the resources are available to develop the needed data. Solving the second problem – the existence of erroneous denials that would not have produced eligible claims if an error had not occurred – could require a redesign of the Denied Claims Accuracy investigation. The third problem is that statistical modeling assumes that erroneously denied claimants are essentially similar to claimants who are found eligible for benefits – in other words, that erroneous denials are random errors. The limited available evidence suggests that this is a valid assumption, although further research will need to be performed to confirm it.

We conclude that, although there are both conceptual and practical problems with estimating the dollar impact of erroneous denials, it may be feasible to overcome the problems and produce estimates of the dollar impact of erroneous denials that could be reasonable guides for policy.

6. Staff Costs of Conducting Investigations

As noted in Chapter 3, the pilot states reported that the staff time needed to investigate an annual case load of 200 monetary denials, 200 separation denials, and 200 nonseparation denials exceeded the 2 staff positions that were allocated to each state. At the Close-Out meeting in November, 1998, the consensus among the states was that 2 staff positions would be consistent with handling an annual case load of 150 of each type of denial, or a total of 450 per year.

Table 6-1 shows the data from pilot states that have made available their estimates of the time spent on investigations during the DCA Pilot Project. Looking first at the estimates of hours spent on each type of investigation, the pilot states appear to fall into two groups. The first comprises Nebraska and West Virginia, both of which estimated that each denied claims investigation took on the order of 4 hours to complete. The second group comprises New Jersey and Wisconsin, both of which estimated that each denied claims investigation took between 8 and 11 hours to complete. (South Carolina's estimates are not yet available.) This is a substantial range, and some discussion of the differences among the pilot states could be illuminating. Because the figures were not available at the time of the November, 1998, Close-Out meeting, these differences were not discussed there. It is possible that, for various reasons that the states could provide, the figures shown in Table 6-1 really are not comparable. It should be noted that, in some cases, the figures in Table 6-1 were provided verbally and are not based on final records maintained

by the pilot states. (The notes to Table 6-1 indicate which of the estimates are verbal and which are final.)

Two states – Nebraska and West Virginia – supplied their figures on total hours spent on the DCA Pilot Project. Nebraska reported approximately 3,700 hours total, and West Virginia reported approximately 4,700 hours total. These are roughly comparable, and differences between the two are likely the result of differences in the amount of supervisory time allocated in the two states.

In addition to the figures displayed in Table 6-1, West Virginia and Wisconsin indicated the total number of staff positions that they believe would be consistent with an annual case load of 600 denied claims investigations. West Virginia estimated that between 2.0 and 2.5 staff positions are needed to perform 600 investigations annually, and an additional 0.5 to 0.75 position needed for support functions. This suggests a total of between 2.5 and 3.25 staff positions. (Note that these estimates appear somewhat high relative to the estimates of hours per investigation shown in Table 6-1; however, those "hours per investigation" estimates are verbal estimates only.) Wisconsin estimated that just over 2.5 staff positions are needed to perform 600 investigations annually, and that 3 positions could be justified once leave time is included. (Wisconsin's estimate of staff positions needed appears consistent with its figures on hours per investigation.)

The above findings suggest that further clarification of the time estimates provided by the pilot states would be useful in determining the staff requirements that are consistent with various case loads.

7. Summary and Recommendations

During fiscal year 1998, UI agencies made over 10.6 million determinations of monetary eligibility for UI benefits, over 3.4 million separation determinations, and over 4.3 million nonseparation determinations (see Chapter 1 for details). The Denied Claims Accuracy Pilot Project was an operational pilot whose purpose was to gather the information that will be needed to guide a program of measuring Denied Claims Accuracy in all 53 UI jurisdictions. In particular, the Unemployment Insurance Service needed answers to the following questions: (a) What are the resource requirements of a Denied Claims Accuracy program? (b) How much does the field-verification approach that has been used to appraise the accuracy of UI payments during the last 10 years (Benefits Quality Control or Benefits Accuracy Measurement) add to the Quality Performance Indicator approach in assessing the accuracy of nonmonetary denials? (c) What are the main reasons for denial errors, and how can these reasons be coded so that all states can identify and eliminate or mitigate denial errors? (d) What are the obstacles to measuring denials errors consistently? (See Chapter 2 for elaboration of these issues.)

7.1. Summary of the Design and Operations of the DCA Pilot Project

In each of five pilot states that had agreed to participate, random samples of (roughly) 200 monetary denials, 200 separation denials, and 200 nonseparation denials

were drawn and subjected to intensive investigation (along the lines of the existing core BQC program) in order to determine their accuracy. Specifically, monetary denials were investigated using the BQC approach, which involved a review of all pertinent agency records, an interview with the claimant, and contacts with base-period employers to ascertain the correct wages, hours of work, weeks of work, etc., as prescribed by State law.

Nonmonetary denials (both separations and nonseparations) were reviewed as follows: (a) The DCA-trained investigator conducted a BQC-type review of the case, involving a claimant interview and appropriate contacts with employers and/or third parties to determine the eligibility decision that would accord with a fully-informed application of State law and policy. (b) A different investigator reviewed the nonmonetary adjudication and all pertinent agency records using the existing Quality Performance Indicator (QPI) instrument. Using that review instrument and applicable agency data alone, this investigator rated each denial according to the seven QPI elements, enabling a point score to be assigned to the case. The BQC supervisor in each pilot state was responsible for reviewing all cases before they were considered final, and the data from investigations were periodically transmitted to the Department of Labor for its own use and for transmission to the contractor. (See section 2.5 for more on the design of the pilot.)

At the beginning of the DCA Pilot Project, a decision was made to provide each state with resources for two DCA Pilot Project investigators. Sampling began in week 36 of 1997 (early September 1997) and continued until between week 33 and week 36 of

1998 (August through early September 1998), depending on the pilot state. Representatives of the National Office, Regional Offices, and the contractor made site visits to four of the five pilot states. The goals of the site visits were to: (a) inform the staffs of the National Office, Regional Offices, and the contractor of the procedures being followed by each state, (b) allow each participating state to advise the National Office, Regional Offices, and the contractor of problems and concerns, (c) help ensure the integrity of the DCA Pilot Project and its outcomes, and (d) ensure the usefulness of the DCA Pilot Project to national implementation of DCA measurement.

The main findings of the site visits include the following:

- The pilot states reported that 3 staff years would be needed to handle a case load of 200 of each type of denial (600 total). Data reported by the states (see Chapter 6) are not fully consistent with this conclusion, but it is the consensus of the pilot states. Sample sizes will have to reflect resource requirements and resource availability.

- Consistent with data summarized in section 4.5, the pilot states reported greater difficulty in obtaining information from claimants during DCA investigations than they experience during core BQC investigations.

- The pilot states reported that difficulty in obtaining responses from claimants caused delays in completing cases. However, the pilot states do favor timeliness standards for the DCA program, recognizing that standards for DCA may need to be structured somewhat differently than existing BQC timeliness standards.

- Except for Wisconsin, the pilot states agree that the sampling of monetary denials should be delayed for two weeks (ten work days) from the date that the claim was filed. Such a delay would avoid including in the samples to be investigated claims that were initially denied but that will be redetermined based on confirmation of additional wages.

- The pilot states reported no greater difficulties with investigations involving claims than with those involving intrastate claims. (Unfortunately, the Data Collection Instrument did not include a code for interstate claims, so a

quantitative analysis of the error rates on denied interstate claims could not be included in Chapter 4.)

- Staffs of the pilot states generally indicated that they prefer intensive investigation of denied claims (as in the DCA Pilot Project) to the Quality Performance Indicator rating of nonmonetary denials. Most feel that the QPI can be misinterpreted as a measure of denied claims accuracy, which it is not. They prefer DCA measurement mainly because it yields a measure of the denied claims accuracy that is what it appears to be.

- The pilot states reported only minor (and easily resolved) start-up problems in implementing the COBOL sampling program and getting the DCA Pilot Project running.

- Most of the pilot states had suggestions for changes in the Data Collection Instrument. Many of the suggested changes pertained to allowing fields in the DCI to be programmed to default (or to be "stamped") to prespecified values for each type of denial investigation. Another frequently noted problem was that the Data Collection Instrument accepts only three digits for a monetary denial error. Finally, some states have specific issues that must be accommodated in the Data Collection Instrument.

- Although the staffs of the pilot states indicated that the transition from the investigation of BQC paid claims to denied claims was fairly smooth, some also believed that training in the BQC methodology in general and in nonmonetary determinations in particular would be beneficial.

See Chapter 3 for further information on the pilot's operation and for the detailed findings from the site visits.

7.2. Summary of Main Findings of the Pilot

Table 7-1 summarizes the essential findings of the 1997-98 DCA Pilot Project regarding the accuracy of denied claims. Column (1) shows the denied claim error rates unadjusted for the effects of appeals, redeterminations, or agency actions to resolve issues. Column (2) shows the error rates adjusted for the effects of appeals and

redeterminations, column (3) shows the error rates adjusted for the effects agency actions to resolve issues, and column (4) shows the error rates adjusted for the effects of appeals, redeterminations, and agency actions to resolve issues.

As discussed in Chapter 4, the incidence of erroneous denials varies substantially, both among monetary, separation, and nonseparation determinations, and among the five pilot states. In the five pilot states together, the unadjusted monetary denial error rate averaged 16 percent, the separation denial error rate averaged 8.7 percent, and the nonseparation denial error rate averaged 15 percent. Both unadjusted and adjusted error rates tend to be lower for separation denials than for monetary and nonseparation denials. Clearly, because of substantial differences among the five pilot states in laws, policies, and procedures, no conclusions about the quality of state administration can or should be drawn from interstate comparisons of the error rates.

Column (2) of Table 7-1 shows that appeals and redeterminations had no impact on monetary denial error rates. Appeals and redeterminations did, however, reduce the separation denial error rate in the five pilot states together by about 2 percentage points (or 22 percent), and reduced the nonseparation denial error rate in the five pilot states by about 1 percentage point (or 6 percent). Appeals and redeterminations, then, are more effective in reversing erroneous nonmonetary denials than in reversing erroneous monetary denials.

Column (3) of Table 7-1 shows that the State UI agencies, through their own actions, reduced the monetary denial error rate in the five pilot states together by nearly

5 percentage points (or nearly 31 percent). The agencies' actions also reduced the separation and nonseparation denial error rates in the five pilot states, but by only about 1 percentage point in each case. Agencies' actions, then, are more effective in reversing erroneous monetary denials than in reversing erroneous nonmonetary denials.

Table 7-1 also displays summary results of the 1987 denied claims pilot. The unadjusted error rates in the 1997-98 DCA Pilot Project are generally lower than those in the earlier pilot (the exception is nonseparation error rates, which are about the same in both pilots). In fact, the unadjusted error rates in the 1997-98 pilot are roughly comparable to the error rates after adjusting for appeals and redeterminations in the earlier pilot (again, with the exception of the nonseparation error rates). This suggests that the determinations process, before any self-correction, has improved over the past decade. However, appeals and redeterminations now do less to reduce denial errors than they did a decade ago. (It appears impossible to adjust the error rates from the 1987 pilot so as to take account of agency resolution of issues. Accordingly, columns (3) and (4) of Table 7-1 have no data on the 1987 pilot.)

Table 7-2 summarizes data on the responsibility for erroneous denials from both the 1997-98 and the 1987 pilots. The figures show that, in 1997-98 in the five pilot states together, the agency alone was responsible for one-third of the monetary denial errors and two-thirds of the nonmonetary denial errors. These percentages are similar to the figures for the 1987 pilot. (As noted in the table, the figures from the 1987 and 1997-98 pilots are not strictly comparable because in 1987, responsibility was assigned to a single party,

whereas in 1997-98, responsibility could be assigned either to a single party or jointly to multiple parties. For the 1997-98 pilot, we also include figures showing the percentage of errors for which the agency was responsible either alone or jointly with another party. Although the agencies appear to have played a role in a higher percentage of monetary denial errors in 1997-98 than in 1987, for nonmonetary denial errors, the percentages remain similar between the two pilots.)

A main objective of the DCA Pilot Project was to compare the results of comprehensive field investigations with the QPI assessment of the quality of the determinations process. Table 7-3 shows cross-tabulations of the accuracy of separation and nonseparation denials by whether the denial determination passed or failed the QPI review. The results suggest that the correlation between QPI and DCA is highly imperfect: In only one case is the coefficient of correlation between the DCA assessment and the QPI assessment greater than 0.3. In all five pilot states and for both separation and nonseparation denials, a significant percentage of denials that were determined improper by DCA passed QPI, and a significant percentage of denials that were determined proper by DCA failed QPI. Moreover, the QPI suggests that the determinations process is less accurate than comprehensive field audits show. The conclusion is that the QPI is a very noisy predictor of the accuracy of denials.

7.3. Recommendations

In addition to the quantitative findings reported in Chapter 4 and briefly summarized

above, the process of visiting states yielded numerous findings that are important to the implementation of the DCA program. Those findings and their implications are summarized here.

7.3.1. Samples Size and Resource Requirements. Each pilot state drew roughly 200 denials of each type during the pilot. At the beginning of the pilot, 2 staff positions were allocated to each pilot state. Several considerations suggested that 2 staff would be able to investigate 600 denials in a year: (a) previous cost studies of time to conduct BAM cases, (b) the assumption that most verifications would be done by telephone, and (c) the assumption that each denial investigation concerned only one issue. Also, based on the previous pilot, error rates were expected to be in the range of 15 to 20 percent.

Denial error rates ranged from somewhat over 20 percent to under 5 percent. These rates mean inferences about the type and cause of error will be based on at most 40 cases with errors in a year. In low error-rate states inferences will be based on 10 or fewer errors in a year. It appears that, once allowance is made for supervisory review, about 2 to 3 staff years, not 2, are needed to investigate 600 denials in a year.

The UI Performs budget currently includes only 2 full-time equivalent positions per state for denials. We recommend maintaining staffing at this level and making the denials annual sample size 150 for each type of denial. Samples of 150 also put the denials total (450) more in line with the total for BAM cases (360 in the 10 smallest states, 480 elsewhere). The smaller samples would be more manageable, especially for smaller states. We would encourage states to gain a year or two of experience with denials

accuracy measurement, and then if they find error rates are particularly high for certain types of denials, adjust sample sizes to focus on the problem area.

Reducing the sample size from 200 cases per year to 150 cases per year will increase the sampling error approximately 15.5 percent. For example, with a 20 percent error rate, the 95 percent confidence interval for a sample of 200 denied claims would be 20% $\pm$ 5.54%. The corresponding interval for a sample of 150 denied claims would be 20% $\pm$ 6.40%. While larger sample sizes are preferable because they yield more precise estimates, the loss of precision in moving from samples of 200 cases to samples of 150 cases is not dramatic. The lower the error rate, the smaller the absolute (but not relative) increase. For example, for an error rate of 10 percent, the 95 percent confidence interval will rise from 10% $\pm$ 4.16% for a sample of 200 cases to 10% $\pm$ 4.80% for a sample of 150 cases.

Reducing the sample size from 200 to 150 cases will also affect the number of cases available for analyses of the cause, responsibility, and point of detection for erroneous denials. A state with a 10 percent error rate would have only 15, instead of 20, cases available for analysis. While a larger subsample of erroneous denials is desirable, even a small number of cases can enable program managers to identify the most common causes and sources of erroneous denials.

As data on denied claims accuracy accumulate, data for several years can be pooled to offer better information on how to improve the program. In addition, after collecting denied claims data for a few years, states could redesign their samples to target

parts of the program that they have identified as sources of errors.

7.3.2. Contacting Denied Claimants. Experience from the first denials pilot indicated that it was more difficult to complete the claimant interview for denied claimants than for paid claimants: many have already returned to work, they see no value in the process unless they believe they were erroneously denied, and the agency cannot threaten to withdraw benefits for failure to report.

Four of the pilot states completed only between 43 and 79 percent of interviews in monetary investigations, and only between 68 and 83 percent of claimant interviews in nonmonetary investigations, although some individual investigators had very high response rates. (Core BQC completion rates are usually 98 percent and above.) The fifth pilot state, South Carolina, completed over 95 percent of its interviews, often by threatening use of its subpoena power (which most State BQC staffs do not have).

Interview completion will always be an inherent difficulty in conducting investigations of denials. When the program is implemented, the experience of states and investigators with the best records should be drawn on and assembled in a handbook to provide advice and techniques for increasing contact rates. Standards for "reasonable attempts and effort" in attempting to contact claimants who have moved, refuse to answer mail, or return telephone calls should also be enunciated, consistent with the BTQ "reasonable attempt" standards.

7.3.3. Sample Exclusions (Interstate Claims). The BQC program has excluded assessment of interstate paid claim accuracy based on the presumed difficulty and

expense of conducting in-person case investigation and verification. With the adoption of telephone and fax verification methodologies, this difficulty should now be eliminated.

In the DCA Pilot Project, states included interstate denials in the sampling frame. They conducted the verifications by phone, mail and fax, the same as they would for intrastate cases they chose to verify in this way. States reported no greater difficulty (or cost) in investigating and assessing the accuracy of interstate denials than intrastate cases. A clear recommendation stemming from the DCA Pilot Project is to include interstate cases in the national program. It should also be feasible to begin drawing interstate cases for paid claims when denied claim accuracy is brought up.

7.3.4. Sample Timing. States' weekly transaction files picked up denials decisions as soon as they were made. These were sampled in the same week as they occurred, and states drew the denials cases and assigned them for investigation as soon as they appeared in the transaction file.

The pilot states found that simultaneous sampling and assignment resulted in the assignment or investigation of some cases that were not true denials because the state did not have time to complete its normal process of determination. The issue occurs mostly with monetary denials where CWC or UCFE wages are outstanding. The state’s initial determination of "inadequate State wages" triggers a denial transaction which may be sampled even though wages are pending. A similar problem is "pro forma" denials in states, such as New Jersey, which require welfare claimants to first establish that they have no UI eligibility.

For the nationwide program, sampling of monetary denials should be delayed for up to 14 days so that wage-request processes are completed and most of the denials in the sample are "true" denials. Nonmonetary denials should be sampled and assigned for investigation at the end of the week in which they occurred, just as in the pilot.

7.3.5. Approach to Case Review. As discussed in section 7.2, the pilot showed almost no significant correlation between the results of the DCA review of denied nonmonetary claims and the QPI review of the quality of nonmonetary decisions. Overall, 26 percent of nonmonetary denials failed the QPI quality review, versus a denial error rate of 15 percent. Although the QPI quality review has some correlation with a DCA assessment, it is at best a "noisy" and erratic indicator of accuracy. The correlation is strongest for cases that passed the nonmonetary quality review: overall, the pilot determined that 9 of 10 denials that passed the QPI quality review reached the correct decision. On the other hand, only one-third of denials that failed the nonmonetary quality review were incorrect. Thus, if states used denials that failed the nonmonetary quality review to identify incorrectly-decided denials, they would be wrong two thirds of the time.

The clear conclusion is that the accuracy of nonmonetary denials is best estimated using the BQC reassessment and reverification method.

7.3.6. Case Review Techniques. The pilot states chose methods of conducting denials investigations during the pilot without restriction. Most preferred to interview claimants in person, although this was generally not feasible due to staffing constraints. As a result, they obtained most other information from a records review or by fax, phone

or mail.

A recommendation stemming from the DCA Pilot Project is that states be given the same flexibility to select among the various review techniques as they have for paid claim accuracy. Since mandated sample sizes and staffing allocations significantly influence state choice of technique, this recommendation also influenced the sample size recommendation. (For example, mandating 200 denials per type per year and providing 2 positions would result in more extensive use of phone-based methods.)

7.3.7. Staff Assignment to Cases. For the DCA Pilot Project, states were urged to designate two investigators for denials and allow the rest of the BQC staff concentrate on BQC cases. However, the pilot states found that the process worked better when all staff shared denials and payment cases. Accordingly, we recommend that states determine the organization structure they wish to employ. The experience of the pilot states will be made available though distribution of the Evaluation Report.

7.3.8. Case Completion Timeliness. The DCA Pilot Project had no standards for timely completion of cases. In most instances of denial, timely completion to ensure accurate recall is not as important with denials as in the case of BQC cases, which involve recollection of whether claimants made work search contacts.

The difficulty that states' staffs experienced in contacting claimants was the main reason case completion lagged. The coordinators recommended that time lapse standards be set to give investigators incentive to complete their efforts timely. Accordingly, we recommend that time lapse standards be established for denials in the nationwide

program, with the understanding that these standards may differ from the BAM standards.

7.3.9. Data Collection Instrument. The pilot was intended to identify data elements which should be captured for denials investigations, or for which definition should be modified. A major question was whether the monetary impact of erroneous denial decisions could be determined. The first pilot had concluded that the dollar impact of erroneous denials could only be estimated using some type of statistical modeling approach.

In the DCA Pilot Project, findings were recorded in a database using a data collection instrument (DCI) modeled on the earlier Denials Pilot DCI and current BQC DCI. In addition, each nonmonetary denial received an independent QPI review using the QPI DCI, the results of which were entered into a separate database (a scaled-down 9056 report) for analysis.

The pilot states identified several elements in the Denied Claims DCI for which edits needed to be modified or which could be pre-filled or "stamped" in case they were inapplicable in the state. The pilot states recommended eliminating all elements designed to capture dollar impacts of denials. States concluded that the data from the dollar impact elements provided little, if any, real useful information. Nebraska staff, in particular, found it extremely difficult to code many dollar impacts because of the way the Nebraska law relates monetary entitlement to nonmonetary eligibility. They concluded that Nebraska's coding was erratic. In addition, they pointed out that the amount of benefits a claimant would ultimately draw depends on a number of factors such as speed of finding work and the existence of other disqualifying conditions, so that making an inference from a weekly

benefit amount or maximum benefit amount is risky.

In light of the above issues and the discussion in Chapter 5, we recommend that denials accuracy be measured only by the percentage of cases decided correctly or incorrectly. If a dollar impact is deemed necessary, additional research would be needed to determine the extent to which generating convincing estimates might be feasible.

Finally, we note that a workgroup of State and Regional staff has met to review all other elements of the Data Collection Instrument to be used in both Denials and Benefits Quality Control and to develop a revised and integrated DCI.